

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 12, 2003, 08:54:59 ; Search time 14 Seconds
(without alignments)
3750.645 Million cell updates/sec

Title: US-09-899-303a-7
Perfect score: 1155
Sequence: 1 ATGTTGGTGAAGTCATCGA.....TGATGATGAACGTGTAATAG 633

Scoring table: BIOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 112892 segs, 41476328 residues

Total number of hits satisfying chosen parameters: 225784

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ n2p.model -DEV=x1p
-O=/cgn2.1/USPTO.spool/US09899303/rnat.12022003.085448.28260/app.query.fasta.1.775
-DB=SwissProt.40 -OPMT=fastan -SUFFIX=n2p.rsp -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNIT=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTPMT=ptc -NRM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09899303.ecgn.1.19.etunal.12022003.085448.28260 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1078	93.3	3010	1	POLG_HCVTM
2	1073	92.9	513	1	POLG_HCVJ2
3	1073	92.9	3010	1	POLG_HCVJA
4	1073	92.9	3010	1	POLG_HCVJT
5	1060	91.8	3010	1	POLG_HCVBK
6	1057	91.5	520	1	POLG_HCVHK
7	1044	90.4	520	1	POLG_HCVH4
8	995	86.1	3011	1	POLG_HCVH
9	992	85.9	3011	1	POLG_HCV1
10	988	85.5	321	1	POLG_HCVTH
11	983	85.1	321	1	POLG_HCVH8
12	978	80.3	309	1	POLG_HCVH7
13	878	76.0	192	1	POLG_HCVH1
14	826	71.5	737	1	POLG_HCVJ1
15	817	70.7	3033	1	POLG_HCVJ8
16	811	70.2	737	1	POLG_HCVJ5
17	802	69.4	3033	1	POLG_HCVJ6
18	155	13.4	138	1	POLG_HCVJ0

C	19	97	8.6	305	1	YT32_STRFR	P20185 streptomyc
C	20	94	8.1	2476	1	ZAN_PIG	Q28983 sus scrofa
C	21	91.5	8.1	483	1	VE2_HPV14	P36783 human papil
C	22	90	8.0	400	1	B3AR_RAT	P26255 rattus norv
C	23	89.5	8.0	408	1	B3AR_HUMAN	P13945 homo sapien
C	24	89.5	7.7	1125	1	MAP4_MOUSE	P27546 mus musculu
C	25	88	7.8	400	1	B3AR_MOUSE	P25962 mus musculu
C	26	87.5	7.8	493	1	VE2_HPV19	P36786 human papil
C	27	87	7.7	2426	1	SON_HUMAN	P18583 homo sapien
C	28	86.5	7.7	1415	1	ICP4_HSVMG	002362 marek's dis
C	29	85	7.6	405	1	B3AR_SHEEP	09x558 ovis aries
C	30	84.5	7.3	821	1	B3AR_RAT	P16259 rattus norv
C	31	84	7.5	358	1	B3AR_FELCA	Q1614 felis silve
C	32	84	7.5	1077	1	AT10_HUMAN	09h324 homo sapien
C	33	84	7.3	1239	1	MEB3_MOUSE	001098 mus musculu
C	34	84	7.3	3014	1	CLRI_HUMAN	09ny66 homo sapien
C	35	83.5	7.2	810	1	CAN3_CHICK	092177 gallus gall
C	36	83	7.4	405	1	B3AR_CANFA	002662 canis fami
C	37	82.5	7.3	342	1	RX_RAT	09j167 rattus norv
C	38	82.5	7.1	821	1	CAN3_MOUSE	064691 mus musculu
C	39	82.5	7.3	827	1	AXNI_RAT	070239 rattus norv
C	40	82.5	7.3	2404	1	SON_MOUSE	09x47 mus musculu
C	41	82	7.1	198	1	VGLX_HSVEL	P25088 equine herp
C	42	82	7.1	383	1	VGLX_HSVK	P32515 equine herp
C	43	82	7.3	418	1	B3AR_MACMU	Q28524 macaca mula
C	44	82	7.3	676	1	ICP0_HSVBJ	P29128 bovine herp
C	45	82	7.1	797	1	VGLX_HSVBE	P28968 equine herp

ALIGNMENTS

RESULT 1
ID POLG_HCVTM STANDARD: PRT: 3010 AA.
AC P29846:
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-APR-1993 (Rel. 41, Last annotation update)
DT 15-JUN-2002 (Rel. 1, Last sequence update)
DE Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
DE Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2
DE (GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
DE (EC 3.4.22.-); Protease/helicase NS3 (P70) (hepativirin)
DE (EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein
DE NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein
DE NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
OS Hepatitis C virus (isolate Taiwan) (HCV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=31645;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-92230206; PubMed-1314449;
RA Chen P.J., Lin M.H., Tai K.F., Liu P.C., Lin C.J., Chen D.S.;
RT "The Taiwanese hepatitis C virus genome: sequence determination and
RT mapping the 5' terminus of viral genomic and antigenomic RNA.";
RL Virology 188:102-113(1992).
CC -1- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
CC HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
CC NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral
CC precursor polyprotein, commonly with Asp or Glu in the P6
CC position, Cys or Thr in P1 and Ser or Ala in P1'.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +
CC (RNA)(N).
CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
CC PROTEIN C AND MRNA.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

FT	CARBOHYD	430	430	N-LINKED	(GLCNAC. . .)	(POTENTIAL)						
FT	CARBOHYD	448	448	N-LINKED <th>(GLCNAC. . .)</th> <th>(POTENTIAL)</th>	(GLCNAC. . .)	(POTENTIAL)						
FT	CARBOHYD	532	532	N-LINKED <th>(GLCNAC. . .)</th> <th>(POTENTIAL)</th>	(GLCNAC. . .)	(POTENTIAL)						
FT	CARBOHYD	556	556	N-LINKED <th>(GLCNAC. . .)</th> <th>(POTENTIAL)</th>	(GLCNAC. . .)	(POTENTIAL)						
FT	CARBOHYD	576	576	N-LINKED <th>(GLCNAC. . .)</th> <th>(POTENTIAL)</th>	(GLCNAC. . .)	(POTENTIAL)						
FT	CARBOHYD	623	623	N-LINKED <th>(GLCNAC. . .)</th> <th>(POTENTIAL)</th>	(GLCNAC. . .)	(POTENTIAL)						
FT	CARBOHYD	645	645	N-LINKED <th>(GLCNAC. . .)</th> <th>(POTENTIAL)</th>	(GLCNAC. . .)	(POTENTIAL)						
FT	CARBOHYD	2041	2041	N-LINKED <th>(GLCNAC. . .)</th> <th>(POTENTIAL)</th>	(GLCNAC. . .)	(POTENTIAL)						
FT	CARBOHYD	2277	2277	N-LINKED <th>(GLCNAC. . .)</th> <th>(POTENTIAL)</th>	(GLCNAC. . .)	(POTENTIAL)						
FT	CARBOHYD	2440	2440	N-LINKED <th>(GLCNAC. . .)</th> <th>(POTENTIAL)</th>	(GLCNAC. . .)	(POTENTIAL)						
FT	CARBOHYD	2788	2788	N-LINKED <th>(GLCNAC. . .)</th> <th>(POTENTIAL)</th>	(GLCNAC. . .)	(POTENTIAL)						
SO	SEQUENCE	3010 AA;	327017 MM;	AA993794F46DB189	CNC64;							
Alignment Scores:												
Pred. No.:	5.76e-89		Length:		3010							
Score:	1073.00		Matches:		198							
Percent Similarity:	98.56%		Conservative:		7							
Best Local Similarity:	95.19%		Mismatches:		3							
Query Match:	92.90%		Indels:		0							
DB:	1		Gaps:		0							
US-09-899-303A-7 (1-633) x POLG_HCVJA (1-3010)												
QY	4	TTGGGTAGGCAATGATACCTTACCTGAGGGGCTGCGCCAGCTCATGGGGTACATTCCG	63									
DB	119	LeuSilyysValIleasPrIleThCysGlyPheAlaAspLeuMetGlyTYrIlePro	138									
QY	64	CTCTCGGCGCCCCCTAGAGGGGGGCTGCCAGAACCCCTGGCGCATGGCGTCCGGGTTCTG	123									
DB	139	LeuValIleYlaIProLeuGlyGlyAlaIleAlaArgAlaLeuAlaHisGlyAlaIValLeu	158									
QY	124	GAGAGCGCGTCAACTATGCAACAGGGAATTGGCTGGTCTTTCTATCTTCTC	183									
DB	159	GluAspIleYalAsnTYrAlaThGlyAsnLeuProGlySerPheSerIlePheLeu	178									
QY	184	TTGGCTTACTGTCCTGTCGACCATTCACAGCTTCGCTTATGAGTGGCCACAGTGTCC	243									
DB	179	LeuAlaLeuLeuSerCysLeuThIleProAlaSerIleATyGlyAlaIArgAsnValSer	198									
QY	244	GGAGTGTACCATGTCACAGAACGACTGCTCCAACTCAAGCAATGTGGTATGAGCAGCGGAC	303									
DB	199	GlyIleTYrHisValThIraAsnAspCysSerAsnSerIleIleValTYrGlyAlaIleAsp	218									
QY	304	ATGATCATGCACACCCCGGGGTGGTGCCCTGCGTGGGAGAACACTTCCCGCTGC	363									
DB	219	MetIleMetHisThIraProGlyCysValProCysValIArgIleAsnPheSerATyGys	238									
QY	364	TGGGTAGCGGTACCCCCAGGCTCGCAGCTTGAAGAACCCAGCGTCCCACTACGACATA	423									
DB	239	TYrAlaIleLeuThIraProThIleAlaIleAlaIArgAsnSerIleProThIraThIle	258									
QY	424	CGAAGCCAGCTCGATGTGTCTGTGGGGGGCGTGGCTTGTGGTCCGTTAGTACGGGGG	483									
DB	259	ArgAlaGlyHisValAspLeuLeuValGlyAlaIleAlaIleuCysSerIleATyGly	278									
QY	484	GATCTGTGGCGATCTGTCTTCTCTGCTCCAGCTGTTCACATCTCGCTGCCCGGAGAT	543									
DB	279	AspLeuCysGlySerValPheLeuValSerGlnLeuPheThIraPheSerProATyGlyTYr	298									
QY	544	GAGACGGTGCAGAGCTCATTTGCTCATGTATTCGCGGCACATACAGGTACACCGTATG	603									
DB	299	GluThIraValIleAspCysAsnCysSerIleTYrProGlyHisValSerGlyHisArgMet	318									
QY	604	GCTTGGGATATGATGATGACTGG	627									
DB	319	AlaTYrAspMetMetMetAsnThIra	326									
RESULT 4												
POLG_HCVJT	STANDARD;		PRT;		3010 AA.							
AC	000269;											
DT	01-APR-1993 (Rel. 25, Created)											
DT	01-APR-1993 (Rel. 25, Last sequence update)											

DT 15-JUN-2002 (Rel. 41, last annotation update)
 DE Genome polypeptide [Contains: Capsid protein C (core protein) (P22);
 DE Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2
 DE (GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
 DE (P21.3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
 DE (P21.3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein
 DE NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein
 DE NS5B (P66) (P70) (RNA-directed RNA polymerase) (P2.7.7.48)].
 OS Hepatitis C virus (isolate HC-J7) (HCV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 NC NCBI_TaxID=31642;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92295714; PubMed=1318627;
 RA Tanaka T., Kato N., Nakagawa M., Ootsuyama Y., Cho M.J.,
 RA Nakazawa T., Hijioka M., Ishimura Y., Shimotohno K.;
 RT Molecular cloning of hepatitis C virus genome from a single Japanese
 RT carrier: sequence variation within the same individual and among
 RT infected individuals.*;
 RL Virus Res. 23:39-53(1992).
 CC -1- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
 CC HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
 CC NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral
 CC precursor polypeptide, commonly with Asp or Glu in the P6
 CC position, Cys or Thr in P1 and Ser or Ala in P1.
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +
 CC (RNA)(N).
 CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
 CC LIPID-ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
 CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
 CC PROTEIN C AND MRNA.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: D11168; BAA01943.1; -
 DR PIR: A45573; A45573.
 DR HSP: P26663; IJXP.
 DR MEROPS: S29.001; -
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR002531; HCV_NS1.
 DR InterPro: IPR002518; HCV_NS2.
 DR InterPro: IPR004109; HCV_NS3.
 DR InterPro: IPR000745; HCV_NS4A.
 DR InterPro: IPR001490; HCV_NS4B.
 DR InterPro: IPR002868; HCV_NS5A.
 DR InterPro: IPR002166; HCV_RdRp.
 DR InterPro: IPR002522; HCV_capsid.
 DR InterPro: IPR002521; HCV_core.
 DR InterPro: IPR002519; HCV_env.
 DR Pfam: PF00938; HCV_RdRp.1.
 DR Pfam: PF01001; HCV_NS4b.1.
 DR Pfam: PF01006; HCV_NS4a.1.
 DR Pfam: PF01506; HCV_NS5a.1.
 DR Pfam: PF01538; HCV_NS2.1.
 DR Pfam: PF01539; HCV_NS2.1.
 DR Pfam: PF01542; HCV_core.1.
 DR Pfam: PF01543; HCV_capsid.1.
 DR Pfam: PF01560; HCV_NS1.1.
 DR Pfam: PF02907; HCV_NS3.1.
 DR ProDom: PD186062; HCV_NS1.1.
 KW Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
 KW Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
 KW Transmembrane; Nonstructural protein; Hydrolase; Serine protease.

FT INT_MET 1 1
 FT CHAIN 1 115
 FT CHAIN 116 191
 FT CHAIN 192 383
 FT CHAIN 384 729
 FT CHAIN 730 1006
 FT CHAIN 1007 1615
 FT CHAIN 1616 1862
 FT CHAIN 1863 2013
 FT CHAIN 2014 3010
 FT TRANSMEM 347 369
 FT ACT_SITE 1083 1083
 FT ACT_SITE 1107 1107
 FT ACT_SITE 1165 1165
 FT NP_BIND 1230 1317
 FT SITE 1313 1319
 FT CARBOHYD 196 196
 FT CARBOHYD 209 209
 FT CARBOHYD 234 234
 FT CARBOHYD 250 250
 FT CARBOHYD 305 305
 FT CARBOHYD 417 417
 FT CARBOHYD 423 423
 FT CARBOHYD 430 430
 FT CARBOHYD 448 448
 FT CARBOHYD 532 532
 FT CARBOHYD 540 540
 FT CARBOHYD 556 556
 FT CARBOHYD 576 576
 FT CARBOHYD 623 623
 FT CARBOHYD 645 645
 FT CARBOHYD 2041 2041
 FT CARBOHYD 2077 2077
 FT CARBOHYD 2240 2240
 FT CARBOHYD 2529 2529
 FT CARBOHYD 2788 2788
 SQ SEQUENCE 3010 AA; 326573 MW; 94A1C77435D642BB CRC64;

 Alignment Scores:
 Pred. No.: 5,76e-89
 Score: 1073.00
 Percent Similarity: 98.08%
 Best Local Similarity: 94.71%
 Query Match: 92.90%
 DB: 1
 Gaps: 0

 US-09-899-303a-7 (1-633) x POLG_HCVUT (1-3010)
 QY 4 TTGGGTAAAGTCATGATACCTTACGTCGGCCGACCTTCATGCGGTACATTCG 63
 DB 119 LeuGlyLysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIlePro 138
 QY 64 CTGTCGGGGCCGCCCTAGGGGGGTGCGCCAGACCCCTGGCGCATGGCGTCCGGTCTG 123
 DB 139 LeuValIleGlyAlaProLeuGlyGlyAlaIleValAlaLeuAlaHisGlyValAlaValIleu 158
 QY 124 GAAGACGGCGGTGAACATAAGCAAGAGGAATTTGCTGGTCTCTTCTATCTTCCTC 183
 DB 159 GluAspGlyValAsnTyrAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeu 178
 QY 184 TTGGCTTACGTCGTCGTCGTCGACATTCGCGCTTATGAGGTCGCGACAGTCC 243
 DB 179 LeuAlaLeuLeuSerCysLeuThrIleProAlaSerAlaTyrGlnValAlaGAsnAlaSer 198
 QY 244 GGGATGATACCATGTCACGACGACGACGACGTCCTCCCAACGACGATGTGTGAGACGCGGAC 303
 DB 199 GlyLeuTyrHisValIleThrAsnAspCysSerAsnSerIleValIleTyrGlnAlaIleGly 218
 QY 304 ATGATCATGACACACCCCGGGTGGTCCCTGCTGGGAGAACATCTTCCGCGTGC 363
 DB 219 MetIleMetHisThrProGlyCysValProCysValAlaGlnAsnAsnAlaSerArgCys 238


```

FT CARBOHYD 417 417 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 423 423 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 430 430 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 532 532 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 540 540 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 556 556 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 576 576 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 623 623 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 645 645 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2041 2041 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2077 2077 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2240 2240 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2529 2529 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2788 2788 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 3010 AA; 327189 MW; F842D5BECFDD9C CRC64;

Alignment Scores:
Pred. No.: 8.59e-88 Length: 3010
Score: 1060.00 Matches: 194
Percent Similarity: 97.12% Conservative: 8
Best Local Similarity: 93.27% Mismatches: 6
Query Match: 91.77% Indels: 0
DB: Gaps: 0

US-09-899-303a-7 (1-633) x POLG_HCVBK (1-3010)
OY 4 TTGGGTAAAGTCATGATACCCCTTACGCGCGCTTGGCGGACGTCATGGGGTACATTCG 63
DB 119 LGGGlylyyvalilleaprrhrlleuthrcyselyphealaspblemeglytyrillepro 138
OY 64 CTCGTGGCGGCCCTAGGGGGTGTGCCAGACCCCTGGCGCATGGCGGTCTGTG 123
DB 139 LGGGlylylalaaprrhrlleuthrcyselyphealaspblemeglytyrillepro 158
OY 124 GAAGACGGCGTGAACCTATGCAACAGGGAATTCCTGCTTCTCTATCTCTC 183
DB 159 GLuaspelylvalasnyrathrlgylasneuproglycysserpheserillepheu 178
OY 184 TTGGCTTACTGTCCTGTCGACCATTCACCTCCGCTTGTGAGGCGGACGTCGCC 243
DB 179 LGGGlylylalaaprrhrlleuthrcyselyphealaspblemeglytyrillepro 198
OY 244 GGGATGACCATGTGCAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGCGAC 303
DB 199 Gyllyleryrhlsvallthrasnaspysersaslaserillevaltyrllglaalalasp 218
OY 304 ATGATCATGCACACCCCGGGTGGCTGCGCTTGGGGAACAACCTCTCCGCTGC 363
DB 219 LGGGlylylalaaprrhrlleuthrcyselyphealaspblemeglytyrillepro 238
OY 364 TTGGGTAAAGTCATGATACCCCTTACGCGCGCTTGGCGGACGTCATGGGGTACATTCG 423
DB 239 LGGGlylylalaaprrhrlleuthrcyselyphealaspblemeglytyrillepro 258
OY 424 CGAGCGCAGCTGATTTGCTGCTGGGCGGCTCTTCTGCTCCGATATGACGCGG 483
DB 259 Argtrghlsvallasprrhrlleuthrcyselyphealaspblemeglytyrillepro 278
OY 484 GATCTTGGCGGATGCTCTCTGCTGCTCCAGCTGTTACACATCTGCTCCGCGGAT 543
DB 279 Asprrhrlsvallasprrhrlleuthrcyselyphealaspblemeglytyrillepro 298
OY 544 GAGACGTCGACGATGATTCATTCATTCGCGGCGGACGTCATGGGGTACATTCG 603
DB 299 Valthrrhrlleuthrcyselyphealaspblemeglytyrillepro 318
OY 604 GCTTGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 627
DB 319 Alatrppasmetmetmetmetmetmetmetmetmetmetmetmetmetmetmetmet 326

RESULT 6
POLG_HCVBK

```

```

ID POLG_HCVBK STANDARD; PRT; 520 AA.
AC 001403;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polypeptide [Contains: Capsid protein C (core protein) (P22);
DE Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2
DE (GP68) (GP70) (NS1)] (Fragment).
OS Hepatitis C virus (isolate HCV-KP) (HCV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=31644;
RN [1]
RP MEDLINE=93019030; PubMed=1383400;
RA Abe K., Inchausti G., Fujisawa K.;
RT "Genomic characterization and mutation rate of hepatitis C virus
RT isolated from a patient who contracted hepatitis during an epidemic
RT of non-A, non-B hepatitis in Japan."
RL J. Gen. Virol. 73:2725-2729(1992).
CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
CC PROTEIN C AND RNA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL: D10687; BAA01529.1; -
DB PIR: J01925; J01925.
DB InterPro: IPR002531; HCV_NSI.
DB InterPro: IPR002522; HCV_capsid.
DB InterPro: IPR002521; HCV_core.
DB InterPro: IPR002519; HCV_env.
DB Pfam: PF01539; HCV_env; 1.
DB Pfam: PF01542; HCV_core; 1.
DB Pfam: PF01543; HCV_capsid; 1.
DB Pfam: PF01560; HCV_NSI; 1.
DB ProDom: PD186062; HCV_NSI; 1.
DB Polypeptide: Glycoprotein: Coat protein; Envelope protein;
DB Transmembrane; Nonstructural protein.
DB INT_MET 1 1
FT FT 1 115 REMOVED FROM CAPSID PROTEIN C BY THE
FT CHAIN 1 115 CELLULAR AMINOPEPTIDASE.
FT CHAIN 116 191 CAPSID PROTEIN C (POTENTIAL).
FT CHAIN 192 383 MATRIX PROTEIN E (POTENTIAL).
FT CHAIN 384 >520 MAJOR ENVELOPE PROTEIN E (POTENTIAL).
FT TRANSMEM 347 369 NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
FT CARBOHYD 196 196 BY SIMILARITY.
FT CARBOHYD 209 209 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 233 233 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 305 305 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 418 418 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 424 424 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 431 431 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 449 449 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT NON_TER 520 520 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 520 AA; 56476 MW; 1D2BD0A6F27349B CRC64;

Alignment Scores:
Pred. No.: 1.24e-87 Length: 520
Score: 1057.00 Matches: 195
Percent Similarity: 96.63% Conservative: 6
Best Local Similarity: 93.75% Mismatches: 7
Query Match: 91.52% Indels: 0
DB: Gaps: 0

```



```

FT CARBOHYD 417 417 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 423 423 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 430 430 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 476 476 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 532 532 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 540 540 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 556 556 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 576 576 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 623 623 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 645 645 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 3011 AA; 327142 MW; 772C8B29CCD94753 CRC64;

```

Alignment Scores:

```

Pred. No.: 6.3e-82 Length: 3011
Score: 995.00 Matches: 178
Percent Similarity: 92.79% Conservative: 15
Best Local Similarity: 85.58% Mismatches: 15
Query Match: 86.15% Indels: 0
DB: 1 Gaps: 0

```

US-09-899-303a-7 (1-633) x POLG_HCVH (1-3011)

```

QY 4 TTGGTAGGTCATGCATACCATCCCTTAGCGGGCTTCCGCGACCCATGGGGTACATTCCG 63
   |||
Db 119 LengllylvalleasphrlleuthrCysglYphenlaaspleuueclYrlllepro 138
QY 64 CTCGTGGGGCCCCCTAGGGGGTGTGCCAGACCCCTGGCGCATGGCGGTCTTG 123
   |||
Db 139 LeuValIGlYAlaProleuglYglYAlaAlaArgAlaLeuAlaHslglYAlaArgValLeu 158
QY 124 GAAGAGCGCGGTGACATGCAACAGGGAATTTCCGTGCTTCTTCTTCTTCTTCTC 133
   |||
Db 159 GlAspGlyAlaAsnlyrAlaThrGlYAsnLeuProGlYcysSerPheSerllePheLeu 178
QY 184 TTGGCTTTACTGTCTGTGACCATTTCCAGCTTCCGCTTATGAGTGGCCAAAGTTC 243
   |||
Db 179 LeuAlaLeuLeuSerCysleuThrValProAlaSerAlaIrrglValAlaGAsnSerSer 198
QY 244 GGGATGTACCATGTACAGAAAGACAGCTCTCCAACTCAAGCATTTGTATGAGGAGCGGAC 303
   |||
Db 199 GlYleuThylHslValThrAsnAspCysProAsnSerSerValValYrGlAlaAlaAsp 218
QY 304 ATGATGATGACACCCCGCGTGGCGTCCGCTGTCGGAGAACAACTTCTCCCGCTGC 353
   |||
Db 219 AlaIleuHslThrProGlYcysValProCysValAlaArglValYAsnAlaSerArgCys 238
QY 364 TTGGTAGGCTCAGCCCAAGCTCGAGCTAGAGAACGCCAGCTCCCACTAGACATA 423
   |||
Db 239 TrpValAlaValThrProThrValAlaThrArgAspGlYcysLeuProThrThrGlnLeu 258
QY 424 CGAGCGCAAGCTGATTTGCTGTGGGGCGCTCTTCTGTCCGCTATGACAGTGGG 483
   |||
Db 259 ArgArgHslIleAspLeuLeuValGlYSerAlaThrLeuCysSerAlaLeuYrValGlY 278
QY 484 GATCTGTGGGAGTCTGTCTTCGCTCCGCTCCCACTGACCACTTCGCTCCCGCGCAT 543
   |||
Db 279 AspLeuCysGlySerValPheLeuValGlYGlnLeuPheThrPheSerProAlaGlnHsls 298
QY 544 GAGACGCTCAGAGCTGCAATTGCTCATCTATCCGCGCACATTAACAGGCTACCGTATG 603
   |||
Db 299 TrpThrThrGlnAspCysAsnGlySerIleYrProGlYHslIleThrGlnHslArgMet 318
QY 604 GCTTGGGATATGATGATGAAGTGG 627
   |||
Db 319 AlaTrpAsnMetMetAsnTrp 326

```

RESULT 9

```

POLG_HCV1 STANDARD: PRT; 3011 AA.
AC P26664;
DT 01-AUG-1992 (rel. 23, Created)
DT 01-AUG-1992 (rel. 23, Last sequence update)

```

```

DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polyprotein [contains: Capsid protein C (Core protein) (P22);
DE Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2
DE (GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
DE (EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepatitisin)
DE (EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein
DE NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein
DE NS5B (P65) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
OS Hepatitis C virus (isolate 1) (HCV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11104;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91172826; PubMed=1848704;
RA Choo Q.-L., Richman K.H., Han J.H., Berger K., Lee C., Dong C.,
RA Gallegos C., Coit D., Medina-Selby A., Barr P.J., Weiner A.J.,
RA Bradley D.W., Kuo G., Houghton M.;
RT "Genetic organization and diversity of the hepatitis C virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:2451-2455(1991).
CC -1- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
CC HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
CC NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral
CC precursor polyprotein, commonly with Asp or Glu in the P6
CC position, Cys or Thr in P1 and Ser or Ala in P1'.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +
CC {RNA}(N).
CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
CC LIPIDPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
CC PROTEIN C AND RNA.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
DR EMBL: M62321; AAA5676.1; -
DR PIR: A39166; GNVG3.
DR HSSP: P27958; IAlV.
DR MEROPS: S29.001; -.
DR MEROPS: U39.001; -.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR002531; HCV_NS1.
DR InterPro: IPR002518; HCV_NS3.
DR InterPro: IPR004109; HCV_NS4.
DR InterPro: IPR000745; HCV_NS4a.
DR InterPro: IPR001490; HCV_NS4b.
DR InterPro: IPR002868; HCV_NS5a.
DR InterPro: IPR002166; HCV_RdRp.
DR InterPro: IPR002522; HCV_capsid.
DR InterPro: IPR002521; HCV_core.
DR InterPro: IPR002519; HCV_env.
DR InterPro: IPR001650; Helicase_C.
DR pfam: PF00271; helicase_C; 1.
DR pfam: PF00998; HCV_RdRp; 1.
DR pfam: PF01001; HCV_NS4b; 1.
DR pfam: PF01006; HCV_NS4a; 1.
DR pfam: PF01506; HCV_NS5a; 1.
DR pfam: PF01538; HCV_NS2; 1.
DR pfam: PF01539; HCV_env; 1.
DR pfam: PF01542; HCV_core; 1.
DR pfam: PF01543; HCV_capsid; 1.
DR pfam: PF01560; HCV_NS1; 1.
DR pfam: PF02907; HCV_NS3; 1.
DR ProDom: PD186062; HCV_NS1; 1.
KW Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
KW Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;

```

KM Transmembrane; Nonstructural protein; Hydrolase; Serine protease.
 FT INT_MET 1 1
 FT CHAIN 1 115
 FT CHAIN 116 191
 FT CHAIN 192 383
 FT CHAIN 384 729
 FT CHAIN 730 1006
 FT CHAIN 1007 1615
 FT CHAIN 1616 1862
 FT CHAIN 1863 2013
 FT CHAIN 2014 3011
 FT TRANSMEM 347 369
 FT ACT_SITE 1083 1083
 FT ACT_SITE 1107 1107
 FT ACT_SITE 1165 1165
 FT NE_BIND 1230 1237
 FT SITE 1316 1319
 FT CARBOHYD 196 196
 FT CARBOHYD 209 209
 FT CARBOHYD 234 234
 FT CARBOHYD 305 305
 FT CARBOHYD 417 417
 FT CARBOHYD 423 423
 FT CARBOHYD 430 430
 FT CARBOHYD 448 448
 FT CARBOHYD 476 476
 FT CARBOHYD 532 532
 FT CARBOHYD 540 540
 FT CARBOHYD 556 556
 FT CARBOHYD 576 576
 FT CARBOHYD 623 623
 FT CARBOHYD 645 645
 FT CARBOHYD 2041 2041
 FT CARBOHYD 2077 2077
 FT CARBOHYD 2240 2240
 FT CARBOHYD 2364 2364
 FT CARBOHYD 2789 2789
 SQ SEQUENCE 3011 AA; 327197 MW; 65F8C9447FCB5AF9 CRC64;

Alignment Scores:
 Pred. No.: 1,18e-81
 Score: 992.00
 Percent Similarity: 92.31%
 Best Local Similarity: 85.58%
 Query Match: 85.89%
 DB: 1
 Gaps: 0

US-09-899-303a-7 (1-633) x POLG_HCV1 (1-3011)

QY 4 TTGGGTAGGTCATGATACCTTACCTGCGGCTTCGCCGACCTCATGGGTACATTCG 63
 DB 119 LcnglylsvalilleasphrleuthrCysglyPhehalaaspLeuMetGlyTyrillePro 138
 QY 64 CTGCTGGGCCCCCCTAGGGGCTGCTCCAGAGCCCTGGCGCATGGCGGCTGCTG 123
 DB 139 LcvalvalGlyalProLeuGlyGlyalalaArgalaleuIahsGlyvalArgvalleu 158
 QY 124 GAGAGCGCGTGAACATGACAGGGAATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 183
 DB 159 GlnspGlyalasnYrIahlnhGlyalasnleuProGlysserPheSerIlePheleu 178
 QY 184 TTGGCTTACTGTCCTGCTGACCATTCAGCTTCCGCTATGAGTGCCGACAGTCTCC 243
 DB 179 LcvalleuLeuSerCysleuthrValProalSerIahYrGlnvalArgasnSerThr 198
 QY 244 GGGATGACCATGTCAGCAAGACGACTGCTCAACTCAAGCATTTGTTAGAGCGGAC 303
 DB 199 GlyleuYrIahsvalTrhPasnspCysPioasSerSerIlevalYrGlnalalaasp 218
 QY 304 ATATGATGACAGCCCGGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363
 DB 219 AlalleuIahsThrProGlyCysvalProCysValArgGlnGlyalasnAlaserArgCys 238

QY 364 TTGGTACGCTACCCCCAGCTGCGAGTAGAAGCCAGCGTCCCACTACGACATA 423
 DB 239 TrpIalalameThrProThrIahValIahTrhArgaspGlylsleuProIahTrhGlnleu 256
 QY 424 CGAGCCGACGTCATTTGCTGCTGGGGCGGCTTTCCTGCTGCTGCTGCTGCTGCTG 483
 DB 259 ArgIahIahsIleasPheleuValGlySerIahTrhdeuCsSerIahleuYrValGly 278
 QY 484 GATCTGGCGGAGTCGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 543
 DB 279 AspleuysGlySerValIahleuValGlyGlnleuThrPheSerProIahArgHis 298
 QY 544 GAGACGGTGCAGAGCAATTCGCTCATCTATCCGCGCCACATACAGGTACACCTATG 603
 DB 299 TrpThrtrhGlnGlyCysasnCsSerIahleuYrProGlyIahsIleTrhGlnIahsArgMet 318
 QY 604 GCTTGGATATGATGATGAACCTG 627
 DB 319 AlatrPasmelMetMetIahsTrp 326

RESULT 10
 POLG_HCVTH STANDARD; PRT; 321 AA.
 ID P27957;
 AC 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Genome polyprotein [contains: Matrix protein (Envelope protein M);
 DE Major envelope protein E; Nonstructural protein NS1 (Fragment).
 OS Hepatitis C virus (isolate TH) (HCV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBL_TaxID=11117;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91112009; PubMed=1846505;
 RA Weiner A.J., Brauer M.J., Rosenblatt J., Richman K.H., Tung J.,
 RA Crawford K., Bonino F., Saracco G., Choo Q.-L., Houghton M.,
 RA Han J.H.;
 RT "Variable and hypervariable domains are found in the regions of HCV
 RT corresponding to the flavivirus envelope and NS1 proteins and the
 RT pestivirus envelope glycoproteins."
 RL Virology 180:842-848(1991).
 CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
 CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
 CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
 CC PROTEIN C AND RNA.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X53134; CAA37294.1; -
 CC InterPro: IPR002531; HCV_NSI.
 CC InterPro: IPR002523; HCV_core.
 CC InterPro: IPR002519; HCV_env.
 CC Pfam: PF01539; HCV_env.1.
 CC Pfam: PF01542; HCV_core.1.
 CC Pfam: PF01560; HCV_NSI.1.
 CC ProDom: PD186062; HCV_NSI.1.
 CC Polyprotein; Glycoprotein; Coat protein; Envelope protein;
 CC Transmembrane; Nonstructural protein.
 KW NON_TER 1
 FT CHAIN <1 75
 FT CHAIN 76 267
 FT CHAIN 268 >321
 FT CARBOHYD 80 80
 FT CARBOHYD 93 93
 MATRIX PROTEIN (POTENTIAL).
 MAJOR ENVELOPE PROTEIN E (POTENTIAL).
 NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_Taxid=11107;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91112009; PubMed=1846505;
 RA Weimer A.J., Brauer M.J., Rosenblatt J., Richman K.H., Tung J.,
 RA Crawford K., Bonito F., Saracco G., Choo Q.-L., Houghton M.,
 RA Han J.H.;
 RT "Variable and hypervariable domains are found in the regions of HCV
 RT corresponding to the flavivirus envelope and NS1 proteins and the
 RT pestivirus envelope glycoproteins.";
 RL Virology 180:842-848(1991).
 CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
 CC LIPIDPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
 CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
 CC PROTEIN C AND MRNA.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X5135; CAA37295.1; -
 DR InterPro: IPR002521; HCV_core.
 DR InterPro: IPR002519; HCV_env.
 DR Pfam: PF01539; HCV_env; 1.
 DR Hepatitis C virus (isolate HC-97) (HCV).
 KM Polypeptide: Glycoprotein; Coat protein; Envelope protein;
 KW Transmembrane.
 FT CHAIN 1 75
 FT CHAIN <1 75
 FT CHAIN 76 >192
 FT CARBOHYD 80 80
 FT CARBOHYD 93 93
 FT CARBOHYD 118 118
 FT CARBOHYD 189 189
 FT NON_TER 192 192
 SQ SEQUENCE 192 AA; 20315 MW; 98E48BF4C35A84C CR664;

 Alignment Scores:
 Pred. No.: 1.53e-71 Length: 192
 Score: 878.00 Matches: 161
 Percent Similarity: 91.05% Conservative: 12
 Best Local Similarity: 84.74% Mismatches: 17
 Query Match: 76.02% Indels: 0
 DB: 1 Gaps: 0

 US-09-899-303a-7 (1-633) x POLG_HCV1 (1-192)
 QY 4 TTGGGTAAGTCATGATGATCCCTTACGTGAGGGCTGCGGACCTCATTGGGATTCG 63
 DB 3 LGGGLyValIleAspHLeuThrCysGlyPheAlaAspLeuGlyTyrIlePro 22
 QY 64 CTCGTGGGCCCCCTAGGGGCTGCGCAGAGCCCTGGCGCATGGCGTGGTTCG 123
 DB 23 LLeuValGlyAlaProLeuGlyGlyAlaAlaArgAlaLeuAlaHisGlyValAlaArgValLeu 42
 QY 124 GAAGAGCGGCTGATGATGACACAGGAAATTCCTGCTTCTTCATCTTCCTC 183
 DB 43 GIUASpGlyValAsnTyrIleThrGlyAsnLeuProGlyCysSerPheSerIlePheLeu 62
 QY 184 TTGGCTTACCTGCTGCTGACCATTCAGCTCCGCTATGAGGTGGGACAGCTG 243
 DB 63 LLeuAlaLeuLeuSerCysLeuThrValProAlaSerAlaTyrGlnValArgAsnSerSer 82
 QY 244 GGGATGACCATGTCACAGACAGACTGCTCAACTGAGCATGTGTATAGCGACGAC 303
 DB 83 GlyLeuTyrHisValThrAsnAspCysProAlaSerSerIleValTyrGlnAlaAlaAsp 102

QY 304 ATGATCATGACACACCCCGGGTGTGCTGCTCGGGAGAACACTTCGCCGCTCC 363
 DB 103 AlaIleLeuHisThrProGlyCysValProCysValIleGlyLysValSerArgCys 122
 QY 364 TGGGTAGCCCTCACCCCGACCTCGACGATAGAACGCCAGCTCCCACTACAGAAAT 423
 DB 123 TrpValAlaMetThrProThrValAlaThrArgAspGlyLysLeuProThrThrGlnLeu 142
 QY 424 CGAGCCACGTCGATTTGCTGCTGGGCGGCTTCTGCTCCGCTATGACGAGGG 483
 DB 143 ArgArgHisIleAspLeuValIleGlySerAlaThrLeuCysSerAlaLeuTyrValGly 162
 QY 484 GATCTGCGGATGCTGCTCTCTGCTCCACAGCTGTACCATCTGCGCTCGCGCAT 543
 DB 163 AspLeuGlyGlySerValPheLeuValIleGlyGlnLeuPheThrPheSerProArgArgHis 182
 QY 544 GAGAGCGTCAGAGCAATGCTCAATC 573
 DB 183 TrpThrThrGlnGlyCysAsnCysSerIle 192

 RESULT 14
 ID POLG_HCV7 STANDARD; PRT; 737 AA.
 AC P27961;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 40, Last sequence update)
 DE 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Genome polypeptide [contains: Capsid protein C (Core protein); Matrix
 DE protein (Envelope protein M); Major envelope protein E; Nonstructural
 DE proteins NS1 and NS2] (fragment).
 OS Hepatitis C virus (isolate HC-97) (HCV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_Taxid=11114;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92230232; PubMed=1314459;
 RA Okamoto H., Kural K., Okada S.I., Yamamoto K., Lizuka H.,
 RA Tanaka T., Fukuda S., Tsuda F., Mishiro S.;
 RT "Full-length sequence of a hepatitis C virus genome having poor
 RT homology to reported isolates: comparative study of four distinct
 RT genotypes";
 RL Virology 188:331-341(1992).
 CC -1- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
 CC HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
 CC NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
 CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
 CC LIPIDPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
 CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
 CC PROTEIN C AND MRNA.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: D10077; BAA00971.1; -
 DR InterPro: IPR002531; HCV_NS1.
 DR InterPro: IPR002522; HCV_capsid.
 DR InterPro: IPR002519; HCV_env.
 DR Pfam: PF01539; HCV_env; 1.
 DR Pfam: PF01542; HCV_core; 1.
 DR Pfam: PF01543; HCV_capsid; 1.
 DR Pfam: PF01560; HCV_NS1; 1.
 DR ProDom: PD186062; HCV_NS1; 1.
 DR Polypeptide: Glycoprotein; Coat protein; Envelope protein;
 KW Transmembrane; Nonstructural protein.
 FT INIT_MET 1 1
 FT REMOVED FROM CAPSID PROTEIN C BY THE
 FT CELLULAR AMINOPEPTIDASE.

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 18, 2003, 05:54:54 ; Search time 46 Seconds
(without alignments)
4220.139 Million cell updates/sec

Title: US-09-899-303A-7

Perfect score: 633
Sequence: 1 ATGTTGGTAAAGTCAATCGA.....TGATGATGACTGTAATAG 633

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/1/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/lna/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/lna/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/lna/PTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/lna/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	633	100.0	633	3	US-08-612-973-7
2	633	100.0	633	4	US-08-927-597-7
3	612.8	96.8	636	3	US-08-612-973-13
4	612.8	96.8	636	4	US-08-927-597-13
5	608.8	96.2	795	3	US-08-612-973-5
6	608.8	96.2	795	4	US-08-927-597-5
7	605.8	95.7	2082	4	US-08-612-973-47
8	605.8	95.7	2082	4	US-08-927-597-47
9	605.8	95.7	2433	3	US-08-612-973-49
10	605.8	95.7	2433	4	US-08-927-597-49
11	561.4	88.7	932	1	US-08-081-072-15
12	561.4	88.7	932	1	US-08-449-093A-15
13	561.4	88.7	1539	2	US-08-470-426B-17
14	561.4	88.7	1863	2	US-08-470-426B-14
15	556.2	87.9	1037	1	US-08-462-195-1
16	556.2	87.9	1037	3	US-08-636-883-1
17	556.2	87.9	1037	3	US-08-127-829-1
18	553.4	87.4	2116	4	US-08-191-160-21
19	553.4	87.4	9599	3	US-09-014-416-4
20	553.4	87.4	9599	3	US-09-014-416-6
21	535.4	84.6	1167	1	US-08-324-977-9
22	535.4	84.6	1167	2	US-08-384-616-9
23	535.4	84.6	1167	2	US-08-904-686A-9
24	535.4	84.6	1167	2	US-09-315-850-9
25	535.4	84.6	1499	1	US-08-324-977-3
26	535.4	84.6	1499	2	US-08-384-616-3
27	535.4	84.6	1499	2	US-08-904-686A-3

28	535.4	84.6	1499	4	US-09-315-850-3	Sequence 3, Appl1
29	535.4	84.6	6039	1	US-08-324-977-11	Sequence 11, Appl1
30	535.4	84.6	6039	2	US-08-384-616-11	Sequence 11, Appl1
31	535.4	84.6	6039	2	US-08-904-686A-11	Sequence 11, Appl1
32	535.4	84.6	6039	4	US-09-315-850-11	Sequence 11, Appl1
33	535.4	84.6	9030	1	US-08-324-977-13	Sequence 13, Appl1
34	535.4	84.6	9030	2	US-08-384-616-13	Sequence 13, Appl1
35	535.4	84.6	9030	2	US-08-904-686A-13	Sequence 13, Appl1
36	535.4	84.6	9030	4	US-09-315-850-13	Sequence 13, Appl1
37	535.4	84.6	9416	2	US-08-324-977-1	Sequence 1, Appl1
38	535.4	84.6	9416	2	US-08-384-616-1	Sequence 1, Appl1
39	535.4	84.6	9416	2	US-08-904-686A-1	Sequence 1, Appl1
40	535.4	84.6	9416	4	US-09-315-850-1	Sequence 1, Appl1
41	535.4	84.6	9416	4	US-08-823-895A-27	Sequence 27, Appl1
42	534.2	84.4	1037	1	US-08-462-195-3	Sequence 3, Appl1
43	534.2	84.4	1037	2	US-08-636-883-3	Sequence 3, Appl1
44	534.2	84.4	1037	3	US-09-127-829-3	Sequence 3, Appl1
45	506.6	80.0	742	1	US-08-081-072-18	Sequence 18, Appl1

ALIGNMENTS

RESULT 1
US-08-612-973-7
Sequence 7, Application US/08612973
Patent No. 6150134
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
APPLICANT: BOSMAN, FONS
APPLICANT: DE MARTYNOFF, GUY
APPLICANT: BUYSE, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612,973
FILING DATE: 11-MAR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 633 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..630
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..627

US-08-612-973-7

Query Match 100.0%; Score 633; DB 3; Length 633;
Best Local Similarity 100.0%; Pred. No. 1,9e-164;
Matches 633; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGTGGGTAAAGTCATGATACCTTACGTCGGCTTCCGACCTCATGGGGTACATT 60
    |||
    1 ATGTGGGTAAAGTCATGATACCTTACGTCGGCTTCCGACCTCATGGGGTACATT 60
QY 61 CCGCTCGTCGGGGCCCCCTAGGGGGTCTGCGACAGCCCTGGGGCATGGGGTTCGGGTT 120
    |||
    61 CCGCTCGTCGGGGCCCCCTAGGGGGTCTGCGACAGCCCTGGGGCATGGGGTTCGGGTT 120
QY 121 CTGGAAGAGCGGCGTGAATATGCAACAGGGAATTTGCCGTGTTCTTTCTATCTTC 180
    |||
    121 CTGGAAGAGCGGCGTGAATATGCAACAGGGAATTTGCCGTGTTCTTTCTATCTTC 180
QY 181 CTCTTGGCTTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
    |||
    181 CTCTTGGCTTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
QY 241 TCCGGGATGTACATGTATGACAGACATGCTCCACTCAGACATGTGTATGAGGACAGG 300
    |||
    241 TCCGGGATGTACATGTATGACAGACATGCTCCACTCAGACATGTGTATGAGGACAGG 300
QY 301 GACATGATCATGACACACCCCGGGTGCCTGCTTCGCTTCGGGAGAACACTCTTCCGC 360
    |||
    301 GACATGATCATGACACACCCCGGGTGCCTGCTTCGCTTCGGGAGAACACTCTTCCGC 360
QY 361 TGCTGGGTAGCGCTACCCCGCTGCTGAGTGAAGAACGCCAGCTCCCACTACGACA 420
    |||
    361 TGCTGGGTAGCGCTACCCCGCTGCTGAGTGAAGAACGCCAGCTCCCACTACGACA 420
QY 421 ATACAGCCCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
    |||
    421 ATACAGCCCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY 481 GGGGATCTCTGGGATCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
    |||
    481 GGGGATCTCTGGGATCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
QY 541 CATGAGAGCGGTGACAGACTGCAATGCTCAATCTATCCCGGACACATACAGTCCAGT 600
    |||
    541 CATGAGAGCGGTGACAGACTGCAATGCTCAATCTATCCCGGACACATACAGTCCAGT 600
QY 601 ATGGCTGGGATATGATGATGAATGTAATAG 633
    |||
    601 ATGGCTGGGATATGATGATGAATGTAATAG 633
Db 601 ATGGCTGGGATATGATGATGAATGTAATAG 633
```

RESULT 2
US-08-927-597-7

Sequence 7, Application US/08927597

Patent No. 6,245,503

GENERAL INFORMATION:

APPLICANT: MAERTENS, GEERT

APPLICANT: BOSMAN, FONS

APPLICANT: DE MARTYNOFF, GUY

APPLICANT: BUYSER, MARIE-ANGE

TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE

NUMBER OF SEQUENCES: 111

CORRESPONDENCE ADDRESS: 111

ADDRESS: NIXON & VANDERHAYE P.C.

STREET: 1100 NORTH GLEBE ROAD

CITY: ARLINGTON

STATE: VIRGINIA

COUNTRY: U.S.A.

ZIP: 22201-4714

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/927,597

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/612,973

FILING DATE: 11-MAR-1996

ATTORNEY/AGENT INFORMATION:

NAME: BYRNE, THOMAS E.

REGISTRATION NUMBER: 32,205

REFERENCE/DOCKET NUMBER: 1487-10

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 816-4100

TELEFAX: (703) 816-4100

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 633 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

MOLECULE TYPE: CDNA

TOPOLOGY: linear

HYPOTHETICAL: NO

ANTI-SENSE: NO

FEATURE:

NAME/KEY: CDS

LOCATION: 1..630

NAME/KEY: mat_peptide

LOCATION: 1..627

US-08-927-597-7

Query Match 100.0%; Score 633; DB 4; Length 633;

Best Local Similarity 100.0%; Pred. No. 1,9e-164;

Matches 633; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGTGGGTAAAGTCATGATACCTTACGTCGGCTTCCGACCTCATGGGGTACATT 60
    |||
    1 ATGTGGGTAAAGTCATGATACCTTACGTCGGCTTCCGACCTCATGGGGTACATT 60
QY 61 CCGCTCGTCGGGGCCCCCTAGGGGGTCTGCGACAGCCCTGGGGCATGGGGTTCGGGTT 120
    |||
    61 CCGCTCGTCGGGGCCCCCTAGGGGGTCTGCGACAGCCCTGGGGCATGGGGTTCGGGTT 120
QY 121 CTGGAAGAGCGGCGTGAATATGCAACAGGGAATTTGCCGTGTTCTTTCTATCTTC 180
    |||
    121 CTGGAAGAGCGGCGTGAATATGCAACAGGGAATTTGCCGTGTTCTTTCTATCTTC 180
QY 181 CTCTTGGCTTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
    |||
    181 CTCTTGGCTTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
QY 241 TCCGGGATGTACATGTATGACAGACATGCTCCACTCAGACATGTGTATGAGGACAGG 300
    |||
    241 TCCGGGATGTACATGTATGACAGACATGCTCCACTCAGACATGTGTATGAGGACAGG 300
QY 301 GACATGATCATGACACACCCCGGGTGCCTGCTTCGCTTCGGGAGAACACTCTTCCGC 360
    |||
    301 GACATGATCATGACACACCCCGGGTGCCTGCTTCGCTTCGGGAGAACACTCTTCCGC 360
QY 361 TGCTGGGTAGCGCTACCCCGCTGCTGAGTGAAGAACGCCAGCTCCCACTACGACA 420
    |||
    361 TGCTGGGTAGCGCTACCCCGCTGCTGAGTGAAGAACGCCAGCTCCCACTACGACA 420
QY 421 ATACAGCCCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
    |||
    421 ATACAGCCCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY 481 GGGGATCTCTGGGATCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
    |||
    481 GGGGATCTCTGGGATCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
Db 481 GGGGATCTCTGGGATCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
```


QY 541 CATGAGAGGTCAGAGTGCATTCATCTATCCCGGCACATAACAGTCAACGT 600
DB 541 CATGAGAGGTCAGAGTGCATTCATCTATCCCGGCACATAACAGTCAACGT 600
QY 601 ATGGCTGGATATGATGATGAATGTAATAG 633
DB 601 ATGGCTGGATATGATGATGAATGTAATAG 633

RESULT 3

US-08-612-973-13
Sequence 13, Application US/08612973
Patent No. 6150134

GENERAL INFORMATION:

APPLICANT: MAERTENS, GEERT
APPLICANT: BOSMAN, FONS
APPLICANT: DE MARTYNOFF, GUY
APPLICANT: BUYSSE, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612,973
FILING DATE: 11-MAR-1996
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:
LENGTH: 636 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..633
FEATURE:
NAME/KEY: mat.peptide
LOCATION: 1..630
US-08-612-973-13

Query Match 96.8%; Score 612.8; DB 3; Length 636;
Best Local Similarity 98.1%; Pred. No. 6,4e-159;
Matches 620; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 ATGTGGGTAAAGTATGATACCTTACGTGCGGCTTCCGACCTATGGGGTACATT 60
DB 1 ATGTGGGTAAAGTATGATACCTTACGTGCGGCTTCCGACCTATGGGGTACATT 60
QY 61 CCGCTGCTGCGGCGCCCTAGGAGGCTCTGCGCAGAGCCCTGCGCATGGCGGT 120
DB 61 CCGCTGCTGCGGCGCCCTAGGAGGCTCTGCGCAGAGCCCTGCGCATGGCGGT 120

QY 121 CTGGAAGAGCGCGTAACTATGCAACAGGAATTTGCTGCTTTCTATATCTTC 180
DB 121 CTGGAAGAGCGCGTAACTATGCAACAGGAATTTGCTGCTTTCTATATCTTC 180
QY 181 CTTTGGGCTTTACGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
DB 181 CTTTGGGCTTTACGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 241 TCCGGATGATACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
DB 241 TCCGGATGATACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
QY 301 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
DB 301 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
QY 361 TGGTGGGTAGAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
DB 361 TGGTGGGTAGAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
QY 421 ATACAGCCGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB 421 ATACAGCCGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 481 GGGATCTCTGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
DB 481 GGGATCTCTGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 541 CATGAGAGGTCAGAGTGCATTCATCTATCCCGGCACATAACAGTCAACGT 600
DB 541 CATGAGAGGTCAGAGTGCATTCATCTATCCCGGCACATAACAGTCAACGT 600
QY 601 ATGGCTGGATATGATGATGAATGTAATAG 632
DB 601 ATGGCTGGATATGATGATGAATGTAATAG 632

RESULT 4

US-08-927-597-13
Sequence 13, Application US/08927597
Patent No. 6245503

GENERAL INFORMATION:

APPLICANT: MAERTENS, GEERT
APPLICANT: BOSMAN, FONS
APPLICANT: DE MARTYNOFF, GUY
APPLICANT: BUYSSE, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/927,597
FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/612,973
FILING DATE: 11-MAR-1996

ATTORNEY/AGENT INFORMATION:

NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 636 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..633
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..630
US-08-927-597-13

Query Match 96.88; Score 612.8; DB 4; Length 636;
Best Local Similarity 98.18; Pred. No. 6.4e-159;
Matches 620; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 ATGTTGGGTAAAGTCATGATGATACCTTACGTCGGGCTTCGCGACCTTCATGGGGTACATT 60
Db 1 ATGCTGGGTAAAGCCATGATGATACCTTACGTCGGGCTTCGCGACCTTCGCGGGTACATT 60
QY 61 CCGCTCGTGGGCGCCCTTAGGGGGTGTCTCCAGAGCCCTGGGCGATGGCGTCCGGGTT 120
Db 61 CCGCTCGTGGGCGCCCTTAGGGGGTGTCTCCAGAGCCCTGGGCGATGGCGTCCGGGTT 120
QY 121 CTGGAAGACGGCGTGAACATATGACAACAGGAATTTGCCGTGCTCTTCTATCTTC 180
Db 121 CTGGAAGACGGCGTGAACATATGACAACAGGAATTTGCCGTGCTCTTCTATCTTC 180
QY 181 CTCCTGGCTTACTGCTCTGCTGACCATTCACCTTCGCTTACGAGGTGCGAAGCTG 240
Db 181 CTCCTGGCTTACTGCTCTGCTGACCATTCACCTTCGCTTACGAGGTGCGAAGCTG 240
QY 241 TCCGGGATGTACATGTACAGAACAGTGTCTCAACTCAAGCATTTGTATGAGGACG 300
Db 241 TCCGGGATGTACATGTACAGAACAGTGTCTCAACTCAAGCATTTGTATGAGGACG 300
QY 301 GACATGATCATGACACACCCCGGGTGGTCCCTGCGTGGGAGAACACTCTTCCGC 360
Db 301 GACATGATCATGACACACCCCGGGTGGTCCCTGCGTGGGAGAACACTCTTCCGC 360
QY 361 TGCCTGGGTAAAGGCTACCCCGAGCTGGCAGCTAGGAACGCGACATCCCACTACGACA 420
Db 361 TGCCTGGGTAAAGGCTACCCCGAGCTGGCAGCTAGGAACGCGACATCCCACTACGACA 420
QY 421 ATACGACGCAAGTCGATTTGCTGTTGGGGGCTGCTTCTGTTCCGCTATAGCTG 480
Db 421 ATACGACGCAAGTCGATTTGCTGTTGGGGGCTGCTTCTGTTCCGCTATAGCTG 480
QY 481 GGGGATCTGCGGATGCTCTTCCTGCTCCAGAGCTGTTCAACATCTGCGCTGCGGG 540
Db 481 GGGGATCTGCGGATGCTCTTCCTGCTCCAGAGCTGTTCAACATCTGCGCTGCGGG 540
QY 541 CATGAGACGGGCGAGCACTGCAATTGTCATATCCCGGCGCAATTAACAGGTACACCT 600
Db 541 CATGAGACGGGCGAGCACTGCAATTGTCATATCCCGGCGCAATTAACAGGTACACCT 600
QY 601 ATGGCTTGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 632
Db 601 ATGGCTTGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 632

RESULT 5
US-08-612-973-5
; Sequence 5, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:

APPLICANT: MAERTENS, GEERT
APPLICANT: BOSMAN, FONS
APPLICANT: DE MARTYNOFF, GUY
APPLICANT: BUYSE, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612,973
FILING DATE: 11-MAR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 795 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..792
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..789
US-08-612-973-5

Query Match 96.28; Score 608.8; DB 3; Length 795;
Best Local Similarity 98.18; Pred. No. 8.6e-158;
Matches 616; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 ATGTTGGGTAAAGTCATGATGATACCTTACGTCGGGCTTCGCGACCTTCATGGGGTACATT 60
Db 1 ATGCTGGGTAAAGCCATGATGATACCTTACGTCGGGCTTCGCGACCTTCGCGGGTACATT 60
QY 61 CCGCTCGTGGGCGCCCTTAGGGGGTGTCTCCAGAGCCCTGGGCGATGGCGTCCGGGTT 120
Db 61 CCGCTCGTGGGCGCCCTTAGGGGGTGTCTCCAGAGCCCTGGGCGATGGCGTCCGGGTT 120
QY 121 CTGGAAGACGGCGTGAACATATGACAACAGGAATTTGCCGTGCTCTTCTATCTTC 180
Db 121 CTGGAAGACGGCGTGAACATATGACAACAGGAATTTGCCGTGCTCTTCTATCTTC 180
QY 181 CTCCTGGCTTACTGCTCTGCTGACCATTCACCTTCGCTTACGAGGTGCGAAGCTG 240
Db 181 CTCCTGGCTTACTGCTCTGCTGACCATTCACCTTCGCTTACGAGGTGCGAAGCTG 240
QY 241 TCCGGGATGTACATGTACAGAACAGTGTCTCAACTCAAGCATTTGTATGAGGACG 300
Db 241 TCCGGGATGTACATGTACAGAACAGTGTCTCAACTCAAGCATTTGTATGAGGACG 300
QY 301 GACATGATCATGACACACCCCGGGTGGTCCCTGCGTGGGAGAACACTCTTCCGC 360
Db 301 GACATGATCATGACACACCCCGGGTGGTCCCTGCGTGGGAGAACACTCTTCCGC 360

```

Db 301 GACATGATCATGCACACCCCGGGTGCCTGCCCTGCCTCGGAGAACACTCTTCCCGC 360
QY 361 TGTGGGTAGGCTACACCCCGACCTGCAGCTAGAACGCCAGCTCCCACTACGACA 420
Db 361 TGTGGGTAGGCTACACCCCGACCTGCAGCTAGAACGCCAGCTCCCACTACGACA 420
QY 421 ATACGAGCCGCGTCGATTTCTCTGTTGGGGCGGCTGCTTCTGTTCCGCTATGACGTG 480
Db 421 ATACGAGCCGCGTCGATTTCTCTGTTGGGGCGGCTGCTTCTGTTCCGCTATGACGTG 480
QY 481 GGGGATCTCTGCGGATGCTCTCTCTGCTCCAGCTGTCCACATCTGCGCTCGCCG 540
Db 481 GGGGATCTCTGCGGATGCTCTCTCTGCTCCAGCTGTCCACATCTGCGCTCGCCG 540
QY 541 CATGAGACGGTGCAGAGCTGCAATTCATCTATCCGCGCACATTAACAGTCAACCGT 600
Db 541 CATGAGACGGTGCAGAGCTGCAATTCATCTATCCGCGCACATTAACAGGTTACCGT 600
QY 601 ATGCTTGGCATATGATGATGAATGCTGT 628
Db 601 ATGCTTGGCATATGATGATGAATGCTGT 628

```

RESULT 6

```

; Sequence 5, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHAYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25 (ERO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 795 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..792

```

FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..789

US-08-927-597-5

Query Match 96.28; Score 608.8; DB 4; Length 795;
Best Local Similarity 98.18; Pred. No. 8.6e-158;
Matches 616; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```

QY 1 ATGTTGGGTAGGTCATGATGATACCTTACGTCGCGGCTTCCGCGACCTTCATGCGGTACATT 60
Db 1 ATGTTGGGTAGGTCATGATGATACCTTACGTCGCGGCTTCCGCGACCTTCATGCGGTACATT 60
QY 61 CCGCTCGTCGCGCGCCCGCTTAGGGGGTCTCCAGAGCCCTCGCGCATAGCGCTCGGGATT 120
Db 61 CCGCTCGTCGCGCGCCCGCTTAGGGGGTCTCCAGAGCCCTCGCGCATAGCGCTCGGGATT 120
QY 121 CTGGAAGACGGCGTGAATATGCAACAGGAAATTTGCTGTTCTCTTCTATATCTTC 180
Db 121 CTGGAAGACGGCGTGAATATGCAACAGGAAATTTGCTGTTCTCTTCTATATCTTC 180
QY 181 CTCTGGCTTACTGTCTCTGCTGACCAATCCAGCTTCCGCTATAGGTGCGCAAGTG 240
Db 181 CTCTGGCTTACTGTCTCTGCTGACCAATCCAGCTTCCGCTATAGGTGCGCAAGTG 240
QY 241 TCCGGATGTACATGTTCACAGACGACTGCTCAACTCAACTCAACTCAACTCAACTCAACT 300
Db 241 TCCGGATGTACATGTTCACAGACGACTGCTCAACTCAACTCAACTCAACTCAACTCAACT 300
QY 301 GACATGATCATGCACACCCCGGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db 301 GACATGATCATGCACACCCCGGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 361 TGTGGGTAGGCTACACCCCGACCTGCGAGTGAAGACGCGACGCTCCCACTACGACA 420
Db 361 TGTGGGTAGGCTACACCCCGACCTGCGAGTGAAGACGCGACGCTCCCACTACGACA 420
QY 421 ATACGAGCCGCGTCGATTTCTCTGTTGGGGCGGCTGCTTCTGTTCCGCTATGACGTG 480
Db 421 ATACGAGCCGCGTCGATTTCTCTGTTGGGGCGGCTGCTTCTGTTCCGCTATGACGTG 480
QY 481 GGGGATCTCTGCGGATGCTCTCTCTGCTCCAGCTGTTCACACATCTGCGCTCGCGG 540
Db 481 GGGGATCTCTGCGGATGCTCTCTCTGCTCCAGCTGTTCACACATCTGCGCTCGCGG 540
QY 541 CATGAGACGGTGCAGAGCTGCAATTCATCTATCCGCGCACATTAACAGTCAACCGT 600
Db 541 CATGAGACGGTGCAGAGCTGCAATTCATCTATCCGCGCACATTAACAGGTTACCGT 600
QY 601 ATGCTTGGCATATGATGATGAATGCTGT 628
Db 601 ATGCTTGGCATATGATGATGAATGCTGT 628

```

RESULT 7

```

; Sequence 47, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHAYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714

```

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612,973
FILING DATE: 11-MAR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 2082 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2079
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..2076
US-08-612-973-47
```

Query Match 95.7%; Score 605.8; DB 3; Length 2082;
Best Local Similarity 98.1%; Pred. No. 7.7e-157;
Matches 613; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```
QY 4 TTGGGTAAGGTCAATCGATTACCTTACGTCGGGCTTCGCCGACCTCATAGGGGTACATTCCG 63
DB 4 TTGGGTAAGGTCAATCGATTACCTTACGTCGGGCTTCGCCGACCTCATAGGGGTACATTCCG 63
QY 64 CTGTGTGGGCCCCCCCCCCTTACGAGGGGGTGTGCGCAGAACCCCTTGGCGCATGGCGTTCG 123
DB 64 CTGTGTGGGCCCCCCCCCCTTACGAGGGGGTGTGCGCAGAACCCCTTGGCGCATGGCGTTCG 123
QY 124 GAAGACGGCGTGAATGCAATGCAAGGGAATTTGGCTGTCTTCTCTATCTTCTC 183
DB 124 GAAGACGGCGTGAATGCAATGCAAGGGAATTTGGCTGTCTTCTCTATCTTCTC 183
QY 184 TTGGCTTGTCTGTCTGTGACCATTCAGCTTCCGCTTATGAGGTGCCAAGCTGTCC 243
DB 184 TTGGCTTGTCTGTCTGTGACCATTCAGCTTCCGCTTATGAGGTGCCAAGCTGTCC 243
QY 244 GCGATGTACCATGTCAAGACGACTGTCTCAACTCAAGCAATTTGTATGAGCAGCGGAC 303
DB 244 GCGATGTACCATGTCAAGACGACTGTCTCAACTCAAGCAATTTGTATGAGCAGCGGAC 303
QY 304 ATGATCATGACACACCCCGGGGTGGTGCCTTGGGTAAGGGAACAATCTTCCCGCTGC 363
DB 304 ATGATCATGACACACCCCGGGGTGGTGCCTTGGGTAAGGGAACAATCTTCCCGCTGC 363
QY 364 TGGGTAGCGCTCAACCCACGCTCGCAGTAGAAGAACCCAGCGTCCCACTAGCAGATA 423
DB 364 TGGGTAGCGCTCAACCCACGCTCGCAGTAGAAGAACCCAGCGTCCCACTAGCAGATA 423
QY 424 CGAGCCGACGTCGATTTGTCTGTGGGGGGGCTGCTTCTTCTCCGTATGATGAGTGGG 483
DB 424 CGAGCCGACGTCGATTTGTCTGTGGGGGGGCTGCTTCTTCTCCGTATGATGAGTGGG 483
QY 484 GATCTGTGGGGAATCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 543
DB 484 GATCTGTGGGGAATCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 543
QY 543 GATCTGTGGGGAATCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 543
DB 543 GATCTGTGGGGAATCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 543
```

```
QY 544 GAGACGGGTGAGAGCTGCAATTCCTCATCTATCCCGCCACATTAACAGGTACCGGTATG 603
DB 544 GAGACGGGTGAGAGCTGCAATTCCTCATCTATCCCGCCACATTAACAGGTACCGGTATG 603
QY 604 GCTTGGATATGATGATGAACTGCT 628
DB 604 GCTTGGATATGATGATGAACTGCT 628
```

RESULT 8

```
US-08-927-597-47
Sequence 47, Application US/08927597
Patent No. 6245503
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
APPLICANT: BOSMAN, FONS
APPLICANT: DE MARTYNOFF, GUY
APPLICANT: BUYS, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/927,597
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/612,973
FILING DATE: 11-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 2082 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2079
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..2076
US-08-927-597-47
```

Query Match 95.7%; Score 605.8; DB 4; Length 2082;
Best Local Similarity 98.1%; Pred. No. 7.7e-157;
Matches 613; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```
QY 4 TTGGGTAAGGTCAATCGATTACCTTACGTCGGGCTTCGCCGACCTCATAGGGGTACATTCCG 63
DB 4 TTGGGTAAGGTCAATCGATTACCTTACGTCGGGCTTCGCCGACCTCATAGGGGTACATTCCG 63
QY 64 CTGTGTGGGCCCCCCCCCCTTACGAGGGGGTGTGCGCAGAACCCCTTGGCGCATGGCGTTCG 123
```

```
Db 64 CTCGTCGGGGCCCCCTAGGGGGGCGTCAGAGGCCCTGGCCATGGCTCGGGGTTCTG 123
Qy 124 GAAGAGGCGCTGAACATGACACAGAGGAATTTGCTGTTCTTCTATCTTCTC 183
Db 124 GAGGACGGCGTGAATATGACACAGAGGAATTTGCCGGTGTCTTCTTCTTCTC 183
Qy 184 TTGGCTTACTGCTCTGCTGTGACCAATTCAGCTTCGCTTATGAGGTGGCAAGCTGTC 243
Db 184 TTGGCTTACTGCTCTGCTGTGACCAATTCAGCTTCGCTTATGAGGTGGCAAGCTGTC 243
Qy 244 GGGATGTCATGTCATGTCACGACAGCTGCTTCAACTCAAGCATTTGTATGAGCAGCGAC 303
Db 244 GGGATGTCATGTCATGTCACGACAGCTGCTTCAACTCAAGCATTTGTATGAGCAGCGAC 303
Qy 304 ATGATCATGTCACACCCCGGGGTGCTGCTGCTTGGGAGAACAACTCTTCCGCTGC 363
Db 304 ATGATCATGTCACACCCCGGGGTGCTGCTGCTTGGGAGAACAACTCTTCCGCTGC 363
Qy 364 TGGTAGCGCTCACCCCGACGCTGCGAGTAGGAACGCCAGCTCCCACTACGACAAATA 423
Db 364 TGGTAGCGCTCACCCCGACGCTGCGAGTAGGAACGCCAGCTCCCACTACGACAAATA 423
Qy 424 CGACCCCACTGTCATTTGCTGCTTGGGGGCTGCTTCTGCTTCCGCTATGTAAGTGGGG 483
Db 424 CGACCCCACTGTCATTTGCTGCTTGGGGGCTGCTTCTGCTTCCGCTATGTAAGTGGGG 483
Qy 484 GATCTCTCGATCTGCTGCTTCTGCTTCCAGCTGTTCCAGCATTCGCGCTCGCGGCAT 543
Db 484 GATCTCTCGATCTGCTGCTTCTGCTTCCAGCTGTTCCAGCATTCGCGCTCGCGGCAT 543
Qy 544 GAGACGGTGCAGAGCTGCAATTCATCTATCCCGGCGACATAACAGTACCGGTANG 603
Db 544 GAGACGGTGCAGAGCTGCAATTCATCTATCCCGGCGACATAACAGTACCGGTANG 603
Qy 604 GCTTGGGATATGATGATGACTGCT 628
Db 604 GCTTGGGATATGATGATGACTGCT 628

RESULT 9
US-08-612-973-49
; Sequence 49, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BOYSE, MARIE-ANGE
; TITLE OF INVENTION: PORIFIED HEPATITIS C VIRUS ENVELOPE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHAYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612.973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
```

```
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2433 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2430
; FEATURE:
; NAME/KEY: mat.peptide
; LOCATION: 1..2427
US-08-612-973-49

Query Match 95.7%; Score 605.8; DB 3; Length 2433;
Best Local Similarity 98.1%; Pred. No. 8e-157;
Matches 613; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 4 TTGGTAGAGTCATGATGATACCTTACAGTGGGCTTGCAGCTTACGATGAGGATATTCG 63
Db 355 TTGGTAGAGTCATGATGATACCTTACAGTGGGCTTGCAGCTTACGATGAGGATATTCG 63
Qy 64 CTCGTCGGGGCCCCCTAGGGGGGCGTCAGAGGCCCTGGCCATGGCTCGGGGTTCTG 123
Db 415 CTCGTCGGGGCCCCCTAGGGGGGCGTCAGAGGCCCTGGCCATGGCTCGGGGTTCTG 474
Qy 124 GAAGAGGCGCTGAACATGACACAGAGGAATTTGCTGTTCTTCTATCTTCTC 183
Db 475 GAGGACGGCGTGAATATGACACAGAGGAATTTGCCGGTGTCTTCTTCTTCTC 534
Qy 184 TTGGCTTACTGCTCTGCTGTGACCAATTCAGCTTCGCTTATGAGGTGGCAAGCTGTC 243
Db 535 TTGGCTTACTGCTCTGCTGTGACCAATTCAGCTTCGCTTATGAGGTGGCAAGCTGTC 594
Qy 244 GGGATGTCATGTCATGTCACGACAGCTGCTTCAACTCAAGCATTTGTATGAGCAGCGAC 303
Db 595 GGGATGTCATGTCATGTCACGACAGCTGCTTCAACTCAAGCATTTGTATGAGCAGCGAC 654
Qy 304 ATGATCATGTCACACCCCGGGGTGCTGCTGCTTGGGAGAACAACTCTTCCGCTGC 363
Db 655 ATGATCATGTCACACCCCGGGGTGCTGCTGCTTGGGAGAACAACTCTTCCGCTGC 714
Qy 364 TGGTAGCGCTCACCCCGACGCTGCGAGTAGGAACGCCAGCTCCCACTACGACAAATA 423
Db 715 TGGTAGCGCTCACCCCGACGCTGCGAGTAGGAACGCCAGCTCCCACTACGACAAATA 774
Qy 424 CGACCCCACTGTCATTTGCTGCTTCCAGCTGTTCCAGCATTCGCGCTCGCGGCAT 543
Db 775 CGACCCCACTGTCATTTGCTGCTTCCAGCTGTTCCAGCATTCGCGCTCGCGGCAT 834
Qy 484 GATCTCTCGATCTGCTGCTTCTGCTTCCAGCTGTTCCAGCATTCGCGCTCGCGGCAT 543
Db 835 GATCTCTCGATCTGCTGCTTCTGCTTCCAGCTGTTCCAGCATTCGCGCTCGCGGCAT 894
Qy 544 GAGACGGTGCAGAGCTGCAATTCATCTATCCCGGCGACATAACAGTACCGGTANG 603
Db 895 GAGACGGTGCAGAGCTGCAATTCATCTATCCCGGCGACATAACAGTACCGGTANG 954
Qy 604 GCTTGGGATATGATGATGACTGCT 628
Db 955 GCTTGGGATATGATGATGACTGCT 979

RESULT 10
US-08-927-597-49
; Sequence 49, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
```

APPLICANT: MARTENS, GEERT
APPLICANT: BOSMAN, FONS
APPLICANT: DE MARTYNOFF, GUY
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHUYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/927,597
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/612,973
FILING DATE: 11-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 2433 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2430
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..2427
US-08-927-597-49

Query Match 95.7%; Score 605.8; DB 4; Length 2433;
Best Local Similarity 98.1%; Pred. No. 8e-157;
Matches 613; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 4 TTGGGTAAAGTCAATGATACCTTACCTGCGGCTTCCGCCACCTCAATGGGTAAATTCG 63
DB 355 TTGGGTAAAGTCAATGATACCTTACCTGCGGCTTCCGCCACCTCAATGGGTAAATTCG 414
QY 64 CTGCTGCGCGCCCCCTAAGGGGGTGTGCTGCGAGAGCCCTGGGGCATGGCGTCCGGGTTCTG 123
DB 415 CTGCTGCGCGCCCCCTAAGGGGGTGTGCTGCGAGAGCCCTGGGGCATGGGGTCTCGGTTCTG 474
QY 124 GAAGACGGCGTGAATGCAACAGGGAATTTGCTGCTGCTTCTTCTATCTTCTC 183
DB 475 GAGAGCGCGGTAATGCAACAGGGAATTTGCGCGGTTCTTCTTCTATCTTCTC 534
QY 184 TTGGCTTAACTGCTCTGCTGACCATTCACGCTTCCGCTTATGAGGTGCGCAAGTGTCC 243
DB 535 TTGGCTTAACTGCTCTGCTGACCATTCACGCTTCCGCTTATGAGGTGCGCAAGTGTCC 594
QY 244 GGGATGACATGTCAGAGCACTGCTCAACTCAAGCAATGCTGATGAGGACGCGAC 303
DB 595 GGGATGACATGTCAGAGCACTGCTCAACTCAAGCAATGCTGATGAGGACGCGAC 654

QY 304 ATGATCATGCACACCCCGGGTGGTCCCTGCGTTGCGGAGAACACTTCCCGCTGC 363
DB 655 ATGATCATGCACACCCCGGGTGGTCCCTGCGTTGCGGAGAACACTTCCCGCTGC 714
QY 364 TGGGTAGCGCTACACCCCGAGCTGCGAGCTAGAGAACCCAGCGTCCCATAGACATA 423
DB 715 TGGGTAGCGCTACACCCCGAGCTGCGAGCTAGAGAACCCAGCGTCCCATAGACATA 774
QY 424 CGAGCCACGTCGATTTGCTGCTGGGGGGCTCTTCTGCTTCCGCTATGATAGTGAGG 483
DB 775 CGAGCCACGTCGATTTGCTGCTGGGGGGCTCTTCTGCTTCCGCTATGATAGTGAGG 834
QY 484 GATCTGCGGATGCTGCTTCTGCTGCTCCAGCTGTTCCACCATCTCGCTCCGCGCAT 543
DB 835 GATCTGCGGATGCTGCTTCTGCTGCTCCAGCTGTTCCACCATCTCGCTCCGCGCAT 894
QY 544 GAGACGCTGCGAGCACTGCAATGCTCAATCTATCTCCGCGCCACATTAACAGTCAACG 603
DB 895 GAGACGCTGCGAGCACTGCAATGCTCAATCTATCTCCGCGCCACATTAACAGTCAACG 954
QY 604 GCTGGGATGATGATGAATGCT 628
DB 955 GCTGGGATGATGATGAATGCT 979

RESULT 11
US-08-081-072-15
Sequence 15, Application US/08081072
Patent No. 5641654
GENERAL INFORMATION:
APPLICANT: NO. 5641654oru MAKI, Kenjiro YAMAGUCHI, Ayumi
APPLICANT: TOYOSHIMA, and Michinori KOHARA
TITLE OF INVENTION: NON-A NON-B HEPATITIS-SPECIFIC
TITLE OF INVENTION: ANTIGEN AND ITS USE IN HEPATITIS
TITLE OF INVENTION: DIAGNOSIS
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: DiKe, Bronstein, Roberts & Cushman
STREET: 130 Water Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-4280
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk, 3.50inch, 1.4MB storage
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh Classic
SOFTWARE: Microsoft Word Version 4.0A
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/081,072
FILING DATE: June 22, 1993
CLASSIFICATION: 435
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 932 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to genomic RNA
US-08-081-072-15

Query Match 88.7%; Score 561.4; DB 1; Length 932;
Best Local Similarity 93.5%; Pred. No. 8.7e-145;
Matches 586; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
QY 4 TTGGGTAAAGTCAATGATACCTTACCTGCGGCTTCCGCCACCTCAATGGGTAAATTCG 63
DB 9 TTGGGTAAAGTCAATGATACCTTACCTGCGGCTTCCGCCACCTCAATGGGTAAATTCG 68
QY 64 CTGCTGCGCGCCCCCTAAGGGGGTGTGCTGCGAGAGCCCTGGGGCATGGCGTCCGGGTTCTG 123
DB 69 CTGCTGCGCGCCCCCTAAGGGGGTGTGCTGCGAGAGCCCTGGGGCATGGGTGTCGCGGTTCTG 128

0Y 124 GAAGACGGCTGAACATATGCAACAGGGAATTTGCTGCTCTTCTATCTTCTC 183
129 GAGAGCGCGGTAATATGCAACAGGGAATTTGCGCGGCTCTCTTCTATCTTCTC 188
0Y 184 TTGGCTTTACTGCTCTCTCTACCATTCACCTTCCTTTATGAGTGGCAGCTGTC 243
189 TTGGCTTTGCTCTCTCTCTACCATTCACCTTCCTTTATGAGTGGCAGCTGTC 248
0Y 244 GGGATGTACCATGTACAGAACGATGCTCCAACTCAAGCATGTGTATGAGGACGGGAC 303
249 GGGATATACCATGTACAGAACGATGCTCCAACTCAAGCATGTGTATGAGGACGGGAC 308
0Y 304 ATGATCATGACACACCCCGGGTGGCTGCTCCCTGCTGGGAGAACACTTCTCCGCTGC 363
309 ATGATCATGACATACCCCGGGTGGCTGCTCCCTGCTGGGAGAACACTTCTCCGCTGC 368
0Y 364 TTGGTACCGCTACACCCCGAGCTCGCAGCTAGAACCGCAGCTCCGACTACGACATA 423
369 TTGGTACCGCTACCTCCAGCTTAGCGGCAAGAACACACCTCCGACTACGACATA 428
0Y 424 CGAGCCAGCTGATGCTGCTGCTGGGGGCTGCTTCTGCTCCGATATGAGTGGG 483
429 CGAGCGATGCTGCTGCTGCTGGGGGCTGCTTCTGCTCCGATATGAGTGGG 488
0Y 484 GATCTGCGGATGCTGCTGCTGCTCCAGCTGCTCCAGCATCTGCTCCGCGGAT 543
489 GATCTGCTGATGCTGCTGCTGCTCCAGCTGCTCCAGCATCTGCTCCGCGGAT 548
0Y 544 GAGACGCTGACGAGCTCAATGCTCAATCTATCCCGCCACATACAGCTACCGCTATG 603
549 GAGACATATACGAGCTCAATGCTCAATCTATCCCGCCACATACAGCTACCGCTATG 608
0Y 604 GCTTGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 630
609 GCTTGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 635
Db

RESULT 12
US-08-449-093A-15
Sequence 15, Application US/08449093A
Patent No. 5662906
GENERAL INFORMATION:
APPLICANT: No. 5662906oru MAKI, Kenjiro YAMAGUCHI, Ayumi
APPLICANT: TOYOSHIMA, and Michinori KOHARA
TITLE OF INVENTION: NON-A NON-B HEPATITIS-SPECIFIC
TITLE OF INVENTION: ANTIGEN AND ITS USE IN HEPATITIS
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dikey, Bronstele, Roberts & Cushman, LLP
STREET: 130 Water Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-4280
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTED Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/449,093A
FILING DATE: May 24, 1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/081,072
FILING DATE: June 22, 1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/726,141
FILING DATE: July 8, 1991
CLASSIFICATION: 424

INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 932 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to genomic RNA
US-08-449-093A-15

Query Match 88.7% Score 561.4; DB 1; Length 932;
Best Local Similarity 93.5% Pred. No. 8,7e-145;
Matches 586; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

0Y 4 TTGGGTAAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 63
Db 9 TTGGGTAAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 68
0Y 64 CTCTGCGGCGCGCGCGCTAGAGGGGTGCTGCGCAGAGCCCTGCGGCAATGCGGGTCTG 123
Db 69 CTCTGCGGCGCGCGCGCTAGAGGGGTGCTGCGCAGAGCCCTGCGGCAATGCGGGTCTG 128
0Y 124 GAAGACGGCTGAACATATGCAACAGGGAATTTGCTGCTCTTCTATCTTCTC 183
129 GAGAGCGCGGTAATATGCAACAGGGAATTTGCGCGGCTCTCTTCTATCTTCTC 188
0Y 184 TTGGCTTTACTGCTCTCTCTACCATTCACCTTCCTTTATGAGTGGCAGCTGTC 243
189 TTGGCTTTGCTCTCTCTCTACCATTCACCTTCCTTTATGAGTGGCAGCTGTC 248
0Y 244 GGGATGTACCATGTACAGAACGATGCTCCAACTCAAGCATGTGTATGAGGACGGGAC 303
249 GGGATATACCATGTACAGAACGATGCTCCAACTCAAGCATGTGTATGAGGACGGGAC 308
0Y 304 ATGATCATGACACACCCCGGGTGGCTGCTCCCTGCTGGGAGAACACTTCTCCGCTGC 363
309 ATGATCATGACATACCCCGGGTGGCTGCTCCCTGCTGGGAGAACACTTCTCCGCTGC 368
0Y 364 TTGGTACCGCTACACCCCGAGCTCGCAGCTAGAACCGCAGCTCCGACTACGACATA 423
369 TTGGTACCGCTACCTCCAGCTTAGCGGCAAGAACACACCTCCGACTACGACATA 428
0Y 424 CGAGCCAGCTGATGCTGCTGCTGGGGGCTGCTTCTGCTCCGATATGAGTGGG 483
429 CGAGCGATGCTGCTGCTGCTGGGGGCTGCTTCTGCTCCGATATGAGTGGG 488
0Y 484 GATCTGCGGATGCTGCTGCTGCTCCAGCTGCTCCAGCATCTGCTCCGCGGAT 543
489 GATCTGCTGATGCTGCTGCTGCTCCAGCTGCTCCAGCATCTGCTCCGCGGAT 548
0Y 544 GAGACGCTGACGAGCTCAATGCTCAATCTATCCCGCCACATACAGCTACCGCTATG 603
549 GAGACATATACGAGCTCAATGCTCAATCTATCCCGCCACATACAGCTACCGCTATG 608
0Y 604 GCTTGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 630
609 GCTTGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 635
Db

RESULT 13
US-08-470-426B-17
Sequence 17, Application US/08470426B
Patent No. 5856458
GENERAL INFORMATION:
APPLICANT: Okamoto, Hiroaki
APPLICANT: Nakamura, Tetsuo
TITLE OF INVENTION: OLIGONUCLEOTIDE PRIMERS, AND THEIR
TITLE OF INVENTION: APPLICATION FOR HIGH-FIDELITY DETECTION OF NON-A, NON-B
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Beveridge, Degrandt, Wellacher & Young,
STREET: L.L.P.
STREET: 1850 M Street, N.W., Suite 800

Db	859	TTGGCTTTGGCTGTCCTGGTTTGGACCAATCCGACGTTCCGTTTGAAGTGGCAACGTTGTC	918
QY	244	GGGATGTAACATGTCACGAAACGACTGTCCTCAATCAAGAATGTGTATATAGGACGAGAC	303
Db	919	GGGATTAATCAATGTCACGAAACGACTGTCCTCAATCAAGAATGTGTATATAGGACGAGAC	978
QY	304	ATGATCATGTCACACCCCGGGTGCGTGCCCTGCGGTTGCGGGAGAACAACTCTTCCGCTGC	363
Db	979	ATGATCATGATPACTCCCGGGTGCGTGCCCTGCGGTTGCGGGAGAACAAAGCTCCCGTTGC	1038
QY	364	TGGGTACGCTCACCCACGCTCGACGCTGAGAACGCGACGCTGCCACTACGACAATA	423
Db	1039	TGGGTACGCTCACTCCCAACGCTCGCGGGCAGGAATCCAGAGGTCCCACTACGACAATA	1098
QY	424	CGAGCCACGCTCAATTTGCTGCTGGGGGGCGGCTTCTGTTCCGCTATGTAACTGGAGG	483
Db	1099	CGAGCCACGCTCACTGCTGCTGGGGGGCGGCTTCTGCTCGCTATGTAACTGGAGG	1158
QY	484	GATCTGTCGGAAATCTGTCCTCTGCTCCACGCTTTCACATCTCGGCTGGCGGGAT	543
Db	1159	GATCTGTCGGGAATCTGTTCTCTCTGCTCCACGCTTTACCTCTTCGCGCTGGCGGGAT	1218
QY	544	GAGACGGTCGAGAGACTGCAATTTGCTCAATCTATCCGGCGACATTAACAGTCACGATG	603
Db	1219	GAGACAGTCGAGGACTGCAACTGCTCAATCTATCCGGCAATTAACAGTCACCGCATG	1278
QY	604	GCTTGGGATATGATGATGAACCTGGTAA	630
Db	1279	GCTTGGGATATGATGATGAACCTGGTAA	1305

```

: TELEPHONE: (703) 413-3000
: TELEFAX: (703) 413-2220
: TELEX: 248855 OPAT UR
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1037 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Hepatitis C virus
: IMMEDIATE SOURCE:
: CLONE: pCC010
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 17..1036
:
: US-08-462-195-1

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: February 12, 2003, 08:55:00 ; Search time 25.5 Seconds
(without alignments)
4772.791 Million cell updates/sec

Title: US-09-899-303a-7
Sequence: 1 ATGTTGGTAAAGCTCATCGA.....TGATGATGAAGCTGATATAG 633

Scoring table:
BLOSUM62
Xgapop 10.0, Ygapop 0.5
Ygapop 10.0, Ygapop 0.5
Fgapop 6.0, Fgapop 7.0
Delop 6.0, Delop 7.0

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 566448

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL-frame+ .n2p.model -DEV-xip
-O/cgrr2.1/USPTO.spool/US09899303.runtat.12022003.085449.28276/app.query.fasta.1.775
-DB-PIR.73 -OPMT-fastan -SUFFIX-n2p.rpr -MINMATCH-0.1 -LOOPEXT-0
-UNITS-bits -START-1 -END-1 -MATRIX-Biosum62 -TRANS-human40.cdi -LIST-45
-DOCALIGN-200 -THR_SCORE-pct -THR_MAX-100 -THR_MIN-0 -ALIGN-15 -MODE-LOCAL
-OUTFMT-pio -NORM-ext -HEADSIZE-500 -MINLEN-0 -MAXLEN-200000000
-USRR-US09899303.ccg.1.1.48 -cunat.12022003.085449.28276 -NCPU-6 -ICPU-3
-NO_XLPEX -NO_MMAR -LARGEROVER -NEG_SCORES-0 -WAIT -LONGLOG -DEV_TIMEOUT-120
-MAIN_TIMEOUT-30 -THREDS-1 -XGAPOP-10 -YGAPEXT-0.5 -Fgapop-6 -Fgapext-7
-YGAPOP-10 -YGAPEXT-0.5 -DELOP-6 -DELEXT-7

Database :

1: PIR.73:*
2: PIR.73:*
3: PIR.73:*
4: PIR.73:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1094	94.7	513	2	PC1284
2	1084	93.9	441	2	S12707
3	1083	93.8	787	2	PN0677
4	1078	93.3	3010	1	GNMWTV
5	1073	92.9	3010	1	GNMWVC
6	1073	92.9	3010	1	A45573
7	1067	92.4	3010	1	S18030
8	1065	92.2	782	2	S18031
9	1063	92.0	782	2	S19876
10	1060	91.8	369	2	S18032
11	1060	91.8	3010	1	GNMWVC
12	1057	91.5	520	2	J01925
13	1057	91.5	520	2	J01925
14	1045	90.5	550	2	JH0711

15	1044	90.4	523	2	J01926	polypeptide - hepa
16	1043	90.3	782	2	S19875	genome polypeptide
17	995	86.1	3011	1	GNMWCH	genome polypeptide
18	992	85.9	3011	1	GNMWVC	genome polypeptide
19	986	85.4	513	2	A44150	structural protein
20	982	85.0	3011	1	S40770	genome polypeptide
21	980	84.8	640	2	J01584	genome polypeptide
22	921	79.7	411	2	JN0265	genome polypeptide
23	891	77.1	411	2	PC2061	genome polypeptide
24	887	76.8	411	2	S41288	genome polypeptide
25	881	76.3	411	2	PC2060	genome polypeptide
26	847	73.3	415	2	PC4407	genome polypeptide
27	843	73.0	3014	1	JC5620	envelope protein -
28	826	71.5	874	2	J00883	genome polypeptide
29	817	70.7	3033	1	GNMWJ8	genome polypeptide
30	802	69.4	874	2	J00881	genome polypeptide
31	802	69.4	3033	1	J01303	genome polypeptide
32	731	63.3	180	2	PC1303	genome polypeptide
33	723	62.6	180	2	PC1304	genome polypeptide
34	721	62.4	180	2	PC1305	genome polypeptide
35	716	62.0	315	2	PS0164	envelope glycoprot
36	699	60.5	315	2	PS0165	envelope glycoprot
37	688	59.6	315	2	PS0165	envelope glycoprot
38	671	58.1	177	2	S32745	genome polypeptide
39	667	57.7	177	2	S32745	genome polypeptide
40	654	56.6	177	2	S32743	genome polypeptide
41	651	56.4	177	2	S32743	genome polypeptide
42	641	55.5	177	2	S32746	genome polypeptide
43	639	55.3	177	2	S32741	genome polypeptide
44	447.5	38.7	193	2	S35630	genome polypeptide
45	365	31.6	189	2	S32740	polypeptide - hepa

ALIGNMENTS

RESULT 1

PC1284 genome polypeptide - hepatitis C virus (isolate HC-J4) (fragment)

C:Species: hepatitis C virus

C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Nov-2000

C:Accession: PC1284

R:Okamoto, H.; Okada, S.; Sugiyama, Y.; Yotsumoto, S.; Tanaka, T.; Yoshizawa, H.; Tsu

Jpn. J. Exp. Med. 60, 167-177, 1990

A:Title: The 5'-terminal sequence of the hepatitis C virus genome.

A:Reference number: PC1284; MUID:9103116; PMID:2170712

A:Accession: PC1284

A:Molecule type: genomic RNA

A:Residues: 1-513 <OK>

A:Cross-references: GB:D00832; NID:g221513; PIDN:BA00706.1; PID:g221514

C:Superfamily: hepatitis C virus genome polypeptide

Alignment Scores:

Pred. No.: 1.63e-92 Length: 513
Score: 1094.00 Matches: 203
Percent Similarity: 99.52% Conservative: 4
Best Local Similarity: 97.60% Mismatches: 1
Query Match: 94.72% Indels: 0
DB: Gaps: 0

US-09-899-303a-7 (1-633) x PC1284 (1-513)

QY	4	TTGGTAAAGCTATGATACCTTACGTCGCGGCTTGCAGCATGATGGGATACATTCG 63
DB	119	LeuValGlyValIleAspThrLeuThcCysGlyPheAlaAspLeuMetGlyTyrIlePro 138
QY	64	CTGCTGGCGCCCGCTAGGGGGTGTCCAGACCCCTGGCGCATGGCGTCCGGGCTCG 123
DB	139	LeuValGlyAlaProLeuGlyGlyAlaAlaArgAlaLeuAlaHisGlyValArgValIleu 158
QY	124	GAAGACGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 183
DB	159	GluAspGlyValAlaThrGlyAlaThrGlyAlaThrGlyAlaThrGlyAlaThrGlyAlaThr 178


```

Db 179 LeuAlaLeuLeuSerCysLeuThrIleProAlaSerAlaTyrGluValArgAsnValSer 198
QY 244 GGGATGACCATGTCACGAGAGACTGCTCAACATCAAGCATTTGTATGAGCAGCGAC 303
    |||:::|||||
Db 199 GtYalYrHisValThrAsnAspCysSerAsnSerIleValTyrGluAlaAlaAsp 218
QY 304 ATGATCATGCACACCCCGGGTGGTGGCCCTGCGCTTGCGGGAACACTTCCCGCTGC 363
    |||:::|||||
Db 219 MetIleMetHisThrProGlyCysValProCysValArgGluAsnAspPheSerArgCys 238
QY 364 TGGGTACCGCGCACCCCGACGCTGCGAGCTAGAACGCGACGCCCGACAGCACAATA 423
    |||:::|||||
Db 239 TrpAlaLeuLeuThrProThrLeuAlaAlaArgAsnSerIleProThrThrIle 258
QY 424 CGAGCCGACGTCGATTTGCTGCTGCGGGCGGCTCTTCTGTCGCTATGACGAGCG 483
    |||:::|||||
Db 259 ArgArgHisIleAspLeuLeuValGtYalAlaAlaIlePheCysSerAlaMetTyrValGtY 278
QY 484 GATCTGCGCGATGCTGCTCTGCTGCTCCAGCTGTTCACCAATCGCGCTGCCGCGCAT 543
    |||:::|||||
Db 279 AspLeuCysGlySerValIlePheLeuValSerGlnLeuPheThrPheSerProArgArgTyr 298
QY 544 GAGCGGTGACGAGCGACATTTGCTCAATGCTATCCGCGCACATACAGCGCACCGTATG 603
    |||:::|||||
Db 299 GluThrValGlnAspCysAsnCysSerIleTyrProGlyHisValSerGlyHisArgMet 318
QY 604 GCTTGGATATGATGATGATGACTGG 627
    |||:::|||||
Db 319 AlaTrpAspMetMetMetAsnTrp 326

```

RESULT 4

GNMVTW
genome polypeptide - hepatitis C virus (strain Taiwan)
N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu

```

C:Species: hepatitis C virus
A:Note: host Homo sapiens (man)
C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
C:Accession: A40244
R:Chem. P.J.; Lin, M.H.; Tai, K.F.; Liu, P.C.; Lin, C.J.; Chen, D.S.
Virology 188, 102-113, 1992
A:Title: The Taiwanese hepatitis C virus genome: sequence determination and mapping the
A:Reference number: A40244; MUID:92230206; PMID:1314449
A:Accession: A40244
A:Molecule type: genomic RNA
A:Residues: 1-3010 <CH>
A:Cross-references: GB:M84754
C:Superfamily: hepatitis C virus genome polypeptide
C:Keywords: ATP; capsid protein C; envelope protein M; glycoprotein; hydrolase; nonstructura
F:1-115/Product: capsid protein C #status predicted <CPC>
F:116-191/Product: envelope protein M #status predicted <EPM>
F:192-389/Product: major envelope protein E #status predicted <NEE>
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F:1007-1615/Product: nonstructural protein NS2 #status predicted <NS2>
F:1230-1237/Product: hepacivirin #status predicted <NS3>
F:1312-1317/Region: nucleotide-binding motif A (P-loop)
F:1316-1319/Region: DEXH motif
F:1616-1862/Product: nonstructural protein NS4 #status predicted <NS4>
F:1863-2013/Product: nonstructural protein NS4 #status predicted <NS4>
F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F:196-209,233,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,207

```

Alignment Scores:

```

Pred. No.: 5,09e-91 Length: 3010
Score: 1078.00 Matches: 200
Percent Similarity: 97.60% Conservative: 3
Best Local Similarity: 96.15% Mismatches: 5
Query Match: 93.33% Indels: 0
DB: Caps: 0

```

US-09-899-303a-7 (1-633) x GNMVTW (1-3010)

```

QY 4 TTGGTAAAGTCATGATACCTTACGTCGCGCTTCGCCGACCTCATGGGATACATTCG 63
    |||:::|||||
Db 119 LeuGtYalYrHisValIleAspThrLeuThrCysGtYalPheAlaAspLeuMetGtYrIlePro 138
QY 64 CTCGTCGGCGCCCGCCAGGGGGTGGTGGCCGAGACCCCGGCGATGGCGTGGCTGCTG 123
    |||:::|||||
Db 139 LeuValGtYalAspLeuGtYalValAlaArgAlaLeuAlaHisGtYalValAlaLeu 158
QY 124 GAAGACGCGCTGACATGACAGGAGAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 183
    |||:::|||||
Db 159 GluAspGtYalAsnTyrAlaThrGtYalAsnLeuProGlyCysSerPheSerIlePheLeu 178
QY 184 TTGGCTTACTGTCCTGCTGACCATTCGACCTTCGCTGCTGCTGCTGCTGCTGCTGCTG 243
    |||:::|||||
Db 179 LeuAlaLeuLeuSerCysLeuThrIleProAlaSerAlaTyrGluValHisAsnValSer 198
QY 244 GGGATGACCATGTCACGAGAGACTGCTCAACATCAAGCATTTGTATGAGCAGCGAC 303
    |||:::|||||
Db 199 GtYalYrHisValThrAsnAspCysSerAsnSerIleValTyrGluAlaAlaAsp 218
QY 304 ATGATCATGCACACCCCGGGTGGTGGCCCTGCGCTTGCGGGAACACTTCCCGCTGC 363
    |||:::|||||
Db 219 MetIleMetHisThrProGlyCysValProCysValArgGluAsnAspPheSerArgCys 238
QY 364 TGGGTACCGCGCACCCCGACGCTGCGAGCTAGAACGCGACGCCCGACAGCACAATA 423
    |||:::|||||
Db 239 TrpAlaLeuLeuThrProThrLeuAlaAlaArgAsnSerValProThrAlaThrIle 258
QY 424 CGAGCCGACGTCGATTTGCTGCTGCGGGCGGCTCTTCTGTCGCTATGACGAGCG 483
    |||:::|||||
Db 259 ArgArgHisIleAspLeuLeuValGtYalAlaAlaIlePheCysSerAlaMetTyrValGtY 278
QY 484 GATCTGCGCGATGCTGCTCTGCTGCTCCAGCTGTTCACCAATCGCGCTGCCGCGCAT 543
    |||:::|||||
Db 279 AspLeuCysGlySerValIlePheLeuValSerGlnLeuPheThrPheSerProArgArgTyr 298
QY 544 GAGCGGTGACGAGCGACATTTGCTCAATGCTATCCGCGCACATACAGCGCACCGTATG 603
    |||:::|||||
Db 299 GluThrValGlnAspCysAsnCysSerIleTyrProGlyHisValThrGlyHisArgMet 318
QY 604 GCTTGGATATGATGATGATGACTGG 627
    |||:::|||||
Db 319 AlaTrpAspMetMetMetAsnTrp 326

```

RESULT 5

GNMVTJ
genome polypeptide - hepatitis C virus (strain J)

```

N:Contains: capsid protein C; envelope protein M; major envelope protein E; nonstruct
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 19-Jan-2001
C:Accession: A39253; PS0086
R:Kato, N.; Hijioka, M.; Ootsuyama, Y.; Nakagawa, M.; Ohkoshi, S.; Sugimura, T.; Shi
Proc. Natl. Acad. Sci. U.S.A. 87, 9524-9528, 1990
A:Title: Molecular cloning of the human hepatitis C virus genome from Japanese patien
A:Reference number: A39253; MUID:91088550; PMID:2175903
A:Accession: A39253
A:Molecule type: genomic RNA
A:Residues: 1-3010 <KAT>
A:Cross-references: GB:D90208; NID:q221610; PIDN:BA14233.1; PID:q221611
R:Kato, N.; Ohkoshi, S.; Shimotohno, K.
Proc. Jpn. Acad. 65B, 219-223, 1989
A:Title: Japanese isolates of the non-A, non-B hepatitis viral genome show sequence v
A:Reference number: PS0085
A:Accession: PS0085
A:Molecule type: genomic RNA
A:Residues: 2650-2707 <KAT>
A:Experimental source: Japanese isolate
C:Comment: The cleavage sites of this polypeptide have not been determined.
C:Superfamily: hepatitis C virus genome polypeptide
C:Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polypeptide; se
F:2-115/Product: capsid protein C #status predicted <CPC>
F:116-191/Product: envelope protein M #status predicted <EPM>

```

F:192-389/Product: major envelope protein E #status predicted <ME>
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F:1007-1615/Product: hepatitis C virus genome from a single Japanese carrier
F:1230-1337/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEXH motif
F:1616-1862/Product: nonstructural protein NS4 #status predicted <NA>
F:1863-2013/Product: nonstructural protein NS5 #status predicted <NA>
F:196-209,234,250,305,325,417,423,430,448,532,556,576,623,645,1213,1255,2041,2077,2240,2

Alignment Scores:

Pred. No.:	1,47e-90	Length:	3010
Score:	1073.00	Matches:	198
Percent Similarity:	98.56%	Conservative:	7
Best Local Similarity:	95.19%	Mismatches:	3
Query Match:	92.90%	Indels:	0
DB:	1	Gaps:	0

US-09-899-303a-7 (1-633) x GMMVCS (1-3010)

QY 4 TTGGGTAAGGTCATGATACCTTACGTCGCGCTTCGCCGACCTCATGGGTCATTCGCG 63
DB 119 LeuGlyLysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTrpIlePro 138
QY 64 CTCGTCGGCGCCCGCTAGGGGGCTGCTCCAGAGCCCTGGCGCATGGCGGTTCG 123
DB 139 LeuValIGlyAlaProLeuGlyGlyAlaAlaArgAlaLeuAlaHisGlyValAlaArgValLeu 158
QY 124 GAAGAGGGGGTGAATATGCAACAGGAAATTCCTGCTTCCTTCTTCTTCTTCTTCTTCTTCT 183
DB 159 GluAspGlyValAsnThrAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeu 178
QY 184 TTGGCTTTACTGTCTGTGTGACCATTCGACCTTCGCTTATGAGTGGCGCAAGCTGCC 243
DB 179 LeuAlaLeuLeuSerCysLeuThrIleProAlaSerAlaTrpIleValAlaArgAsnValSer 198
QY 244 GGGATGACCATGTCACGACGACTGCTCCAACTCAAGCATTTGTATGAGCGAGCGGAC 303
DB 199 GlyIleThrHisValThrAsnAspCysSerAsnSerSerIleValIleGluAlaAlaAsp 218
QY 304 ATGATCATGCACACCGCGGGTGGCGCTGCGCTTCGGGAGAACAACTTCTCCCGCTGC 363
DB 219 MetIleMetHisThrProGlyCysValProCysValArgGlnSerAsnPheSerArgCys 238
QY 364 TTGGAGCGCTCACCCCAAGCTGCGAGCTAGGAGCAAGCGCTCCCACTAGACAAATA 423
DB 239 TrpValAlaLeuThrProThrIleuAlaAlaArgAsnSerSerIleProThrThrIle 258
QY 424 CGAGCCGACGTCGATTTGCTCGTTGGGGGGCTGCTTCTGTTCCGCTATGACGTGGG 483
DB 259 ArgArgHisValAspLeuLeuValIGlyAlaAlaAlaLeuCysSerAlaMetIleValIGly 278
QY 484 GATCTGAGGAGATCTGCTTCCTGCTGCTCCGAGCTGTCACCATTCGCTCCGCGCGCAT 543
DB 279 AspLeuCysGlySerValPheLeuValSerGlnLeuPheThrPheSerProAlaGlyTrp 298
QY 544 GAGAGCGTCAGAGCTCAATTCATATCCCGGACATCAACAGCTCCACGCTATG 603
DB 299 GluThrValGlnAspCysAsnSerIleTrpProGlyHisValSerGlyHisAlaGlyMet 318
QY 604 GCTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 627
DB 319 AlaTrpAspMetMetMetAsnTrp 326

RESULT 6

genome polypeptide - hepatitis C virus (strain J1)
NA:contains: capsid protein C; envelope protein M; hepatitis C virus (EC 3.4.21.98) (nonstructu
C:Species: hepatitis C virus
C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 19-Jan-2001

C:Accession: A45573
R:Ranka, T.; Kato, N.; Nakagawa, M.; Ootsuyama, Y.; Cht, M.J.; Nakazawa, T.; HijiKat
Virus Res. 23, 39-53, 1992
A:Title: Molecular cloning of hepatitis C virus genome from a single Japanese carrier
A:Reference number: A45573; MUID:92295714; PMID:1318627
A:Accession: A45573
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-3010 <TAN>
A:Cross-references: GB:DI1168; GB:DO1171; NID:9221612; PIDN:BA01943.1; PID:9221613
A:Experimental source: HCV-JT
A:Note: sequence extracted from NCBI backbone (NCBI:106206, NCBI:106207)
C:Superfamily: hepatitis C virus genome polypeptide
C:Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polypeptide; se
F:2-115/Product: capsid protein C #status predicted <CP>
F:116-191/Product: envelope protein M #status predicted <EMP>
F:192-389/Product: major envelope protein E #status predicted <ME>
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F:1007-1615/Product: hepatitis C virus genome polypeptide
F:1230-1337/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEXH motif
F:1616-1862/Product: nonstructural protein NS4 #status predicted <NA>
F:1863-2013/Product: nonstructural protein NS5 #status predicted <NA>
F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS>
Alignment Scores:
Pred. No.: 1,47e-90 Length: 3010
Score: 1073.00 Matches: 197
Percent Similarity: 98.08% Conservative: 7
Best Local Similarity: 94.71% Mismatches: 4
Query Match: 92.90% Indels: 0
DB: 1 Gaps: 0
US-09-899-303a-7 (1-633) x A45573 (1-3010)
QY 4 TTGGGTAAGGTCATGATACCTTACGTCGCGCTTCGCCGACCTCATGGGTCATTCGCG 63
DB 119 LeuGlyLysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTrpIlePro 138
QY 64 CTCGTCGGCGCCCGCTAGGGGGCTGCTCCAGAGCCCTGGCGCATGGCGGTTCG 123
DB 139 LeuValIGlyAlaProLeuGlyGlyAlaAlaArgAlaLeuAlaHisGlyValAlaArgValLeu 158
QY 124 GAAGAGGGGGTGAATATGCAACAGGAAATTCCTGCTTCCTTCTTCTTCTTCTTCTTCTTCT 183
DB 159 GluAspGlyValAsnThrAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeu 178
QY 184 TTGGCTTTACTGTCTGTGTGACCATTCGACCTTCGCTTATGAGTGGCGCAAGCTGCC 243
DB 179 LeuAlaLeuLeuSerCysLeuThrIleProAlaSerAlaTrpIleValAlaArgAsnValSer 198
QY 244 GGGATGACCATGTCACGACGACTGCTCCAACTCAAGCATTTGTATGAGCGAGCGGAC 303
DB 199 GlyIleThrHisValThrAsnAspCysSerAsnSerSerIleValIleGluAlaAlaAsp 218
QY 304 ATGATCATGCACACCGCGGGTGGCGCTGCGCTTCGGGAGAACAACTTCTCCCGCTGC 363
DB 219 MetIleMetHisThrProGlyCysValProCysValArgGlnSerAsnPheSerArgCys 238
QY 364 TTGGAGCGCTCACCCCAAGCTGCGAGCTAGGAGCAAGCGCTCCCACTAGACAAATA 423
DB 239 TrpValAlaLeuThrProThrIleuAlaAlaArgAsnSerSerIleProThrThrIle 258
QY 424 CGAGCCGACGTCGATTTGCTCGTTGGGGGGCTGCTTCTGTTCCGCTATGACGTGGG 483
DB 259 ArgArgHisValAspLeuLeuValIGlyAlaAlaAlaLeuCysSerAlaMetIleValIGly 278
QY 484 GATCTGAGGAGATCTGCTTCCTGCTGCTCCGAGCTGTCACCATTCGCTCCGCGCGCAT 543
DB 279 AspLeuCysGlySerValPheLeuValSerGlnLeuPheThrPheSerProAlaGlyTrp 298

QY 544 GAGACGGTGCAGACTGCAATTCGTCATCTCTCCCGGCCACATTAACAGTACCCTATG 603
|||||
Db 299 GTATTGValGlnAAsPCysAnscysSerIleTyProGIyHIsAlstercIyHisArgMet 318
|||||

QY 604 GCTTGGAATGATGATGAACTGG 627
|||||
Db 319 AAtTPPaSPmeMeMeMetcAnsTrP 326
|||||

RESULT 7
S18030
genome polypeptide - hepatitis C virus (isolate JKI)
N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstruct
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
A.Variety: isolate JKI
C.Date: 19-May-2000 #sequence-revision 19-May-2000 #text-change 23-Mar-2001
C.Accession: S18030; S33570; A48332; S18029
R.Honda, M.; Kaneo, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
submitted to the EMBL Data Library, September 1991
A.Description: A whole genome of hepatitis C virus cDNA was isolated from a single patie
A.Reference number: S18028
A.Accession: S18030
A.Molecule type: genomic RNA
A.Residues: 1-3010 <HON>
A.Cross-references: EMBL.X61596; NID:g59478; PIDN:CAA43793.1; PID:g59479
A.Experimental source: Isolate JKI from an individual
R.Honda, M.; Kaneo, S.; Uncoure, M.; Kobayashi, K.; Murakami, S.
Arch. Virol. 128, 163-169, 1993
A.Title: Sequence analysis of putative structural regions of hepatitis C virus isolated
A.Reference number: A48332; MUID:93119270; PMID:8380322
A.Accession: S33570
A.Molecule type: genomic RNA
A.Residues: 1-547,'T','549-621','V','623-624','S','626-652','DL','655-761','T','763-782<HOW>
A.Cross-references: EMBL.X61591
A.Note: this sequence is inconsistent with the nucleotide translation
as Trp, and Trp for residue 771 as Ser
A.Note: sequence extracted from NCBI backbone (NCBIN:121747, NCBIP:121748)
C.Superfamily: hepatitis C virus genome polypeptide
C.Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polypeptide; serine
F.2-115/Product: capsid protein C #status predicted <CP>
F.116-181/Product: envelope protein M #status predicted <EPM>
F.192-389/Product: major envelope protein E #status predicted <MEP>
F.390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F.730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F.1007-1615/Product: hepacivirin #status predicted <NS3>
F.1230-1237/Region: nucleotide-binding motif A (P-loop)
F.1312-1317/Region: nucleotide-binding motif B
F.1316-1319/Region: DXH motif
F.1616-1862/Product: nonstructural protein NS4a #status predicted <NS4a>
F.1863-2013/Product: nonstructural protein NS4b #status predicted <NS4b>
F.2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F.196,209,224,250,305,417,423,448,532,540,556,576,623,645/Binding site: carbohydrate (A

Alignment Scores:
Pred. No.: 5,24e-90 Length: 3010
Score: 1067.00 Matches: 195
Percent Similarity: 97.60% Conservative: 8
Best Local Similarity: 93.75% Mismatches: 5
Query Match: 92.38% Indels: 0
DB: 1 Gaps: 0

US-09-899-303A-7 (1-633) x S18030 (1-3010)

QY 4 TTGGTAGGTGATCATGATACCTTACGTGCGGCTTCGCGACCTCATGGGTACATTCGG 63
|||||
Db 119 LeuGLyLysValIleAspThrLeuThyrGlyPheAlaAspLeuMetIlyrIlePro 138
|||||

QY 64 CTCGTGCGGCCCCCTAGGGGGGTCTGCGACAGAGCCCTGGCCATGAGGCTCGGGTTCTG 123
|||||
Db 139 LeuValIGLyAlaProLeuGlylAlaAlaArgAlaLeuAlaHisIGLyAlaArgValLeu 158
|||||

QY 124 GAAGACGGCGTGAACATATGCAACAGGGAATTTGCTGTCTCTTCTATCTTCCTC 183

DB 159 GUAAGGUYVALASTPTTALATHRGILYASINLEUPROGLYCYSERPHESEITLPHENEU 178
DB 184 TTGGCTTAATCTGTCTGTCTGCACATTCCAGCTTCGCCGTTTNGAGTGCGGACGTGCC 243
DB 179 LEUALALEULEUSERCYLEUTHETHVALPROVALSERITHTTYTGILVALAIAVSVALSER 198
DB 244 GGAGTAGTACCATGTCACAGAACGACTGCTCCAATCAACCATTTGTATAGGACAGCGAC 303
DB 199 GLYVALITYHISVALITHRASNAPCYSERIASNSERTILEVALTYGLUALAIALASP 218
DB 304 ATGATCATGACACACCCCCGGGTGGCGTCCCTGCGTTCCGGAGAACAACTTTCCCGCTGC 363
DB 219 METIEMETIHSTHPROGLICYSVAILPROCVSVALAIGUGILYASINSERSERARGCS 238
DB 364 TGGGATGAGGCTCACCCCACGCTGCACCTAGTGAAGAAGCAGAGGTCGCCCATACGACAATA 423
DB 239 TRPVALLALEUTHTRPOTHLEUALAIALARGANSERITLIEPTHTHILLIE 258
DB 424 CGACGCCACATCGATTTGCTGTTGGGGCGCGCTGTTCTGTCCGCTATGATAGCTGGGG 483
DB 259 ARGAGHSIVALASPLEULENVALGILYALALAIALAUCYSERIALAMETTYVALGILY 278
DB 484 GATCTCTCGGATCTGCTCTCTGCTCTCCACAGCTGTTCACCAATGCCCTCCGGCGAT 543
DB 279 ASPLKCYSGLYSERVALPHELEVALSERGLNEUPHERTPHESERPROARGATGYR 298
DB 544 GAGACGGTGCAGAGGATGGAATTTGCTCAATCTATATCCGGGCAATACAGAGTACCAGTATG 603
DB 299 GIUTRVAILGHNASCYASMSYSSERLEUTHYTRPROGLYNHISVALSERGLYNHISARGMET 318
DB 604 GCCTTGATATGATGATGATGATGATG 627
DB 319 ALATRPASPMELMELMELASNTIP 326

RESULT 8
SI18031
genome polypeptide - hepatitis C virus (isolate JK2) (fragment)
N.Contains: core protein; envelope protein 1; nonstructural protein 2; NS1/E2 proteins
C.Species: hepatitis C virus
A.Variety: isolate JK2
C.Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 17-Nov-2000
C.Accession: SI18031
R.Honda, M.; Kameo, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
submitted to the EMBL Data Library, September 1991
A.Description: Sequence analysis of putative structural regions of Hepatitis C Virus
A.Reference number: SI18029
A.Accession: SI18031
A.Molecule type: genomic RNA
A.Residues: 1-782 <HON>
A.Cross-references: EMBL:X61593
A.Experimental source: isolate JK2
C.Superfamily: hepatitis C virus genome polypeptide
C.Keywords: capsid protein; core protein; envelope protein; glycoprotein; nonstructural
F.1-191/Product: core protein #status predicted <MAT1>
F.192-383/Product: envelope protein 1 #status predicted <MAT2>
F.384-733/Product: NS1/E2 protein #status predicted <MAT3>
F.734-782/Product: nonstructural protein 2 (fragment) #status predicted <MAT4>

Alignment Scores:
Pred. No.: 7.7e-90 Length: 782
Score: 1065.00 Matches: 196
Percent Similarity: 98.08% Conservative: 8
Best Local Similarity: 94.23% Mismatches: 4
Query Match: 92.21% Indels: 0
DB: 2 Gaps: 0

US-09-899-303A-7 (1-633) x SI18031 (1-782)

DQ 4 TTGGTATAGCATGATGATACCTTACGTCGCGCTTCGCGCACTCATGGGTACATTCG 63
DB 119 LEUGLYLVSVAILLEASPTHTREUTHRCYSGLYPEALASPLEUMETGLTYTTLLEPRO 138

```

QY 64 CTGTCGGGCCCCCCTAGGGGGTGCCTCCAGAGCCCTGGCGCATGGCTCCGGTCTG 123
    |||
Db 139 LeuValGlyAlaProLeuGlyGlyAlaAlaArgAlaLeuAlaHisGlyValAlaGValLeu 158
QY 124 GAAGAGCGCGCTGAACATGACAGAGGAATTCCTGCTGCTGCTTCCTATCTTCCTC 183
    |||
Db 159 GluAspGlyValAsnIleThrAlaThrGlyAsnLeuProGlySerPheSerIlePheLeu 178
QY 184 TTGGCTTTACTGCTGCTGTGACCAATTCAGCTTCGCTTATGAGTGGCGCAACGTGCC 243
    |||
Db 179 LeuAlaLeuLeuSerCysLeuThrIleProAlaSerAlaTyrGluValAlaGValSer 198
QY 244 GGGATGACCATGTCAGAGAGAGCTGCTCACTCAAGCAATTCGTATGAGGAGCGGAC 303
    |||
Db 199 GlyLeuThrHisValThrAsnAspCysSerAsnSerSerIleValIleGluAlaAlaAsp 218
QY 304 ATGATCATGACACACCCCGGGTGCCTGCGCTTCGCGGAGAAACAATCTCCGCTGC 363
    |||
Db 219 LeuIleMetHisThrProGlyCysValProCysValLeuGluIleAsnSerSerArgCys 238
QY 364 TTGGTAGCGCTCAACCCCGGCTGCGCTGCGCTTCGCGGAGAAACAATCTCCGCTGC 423
    |||
Db 239 TrpValAlaLeuThrProThrLeuAlaAlaArgAsnSerSerIleProThrThrIle 258
QY 424 CGAGCGCAGCTGCATTCCTGCTGCTGCGGCGGCTTCCTGCTTCGCTATGACGTGGG 483
    |||
Db 259 ArgArgHisValAspLeuLeuValIleGlyAlaAlaAlaPheCysSerAlaMetIleValGly 278
QY 484 GATCTGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 543
    |||
Db 279 AspLeuCysGlySerValAlaPheLeuValSerGlnLeuPheThrPheSerProAlaGlyTyr 298
QY 544 GAGACGCTGACAGAGTCGCAATTCCTCAATCTATCCGGCCATCAACAGGTCACCGTATG 603
    |||
Db 299 GluThrValGlnAspCysAsnCysSerLeuTyrSerGlyHisValSerGlyHisArgMet 318
QY 604 GCTTGGATATGATGATGAACCTGG 627
    |||
Db 319 AlaTrpAspMetMetMetAsnTrp 326

```

RESULT 9

S19876
 genome polypeptin - hepatitis C virus (isolate JK5) (fragment)
 N:contains: core protein; envelope protein 1; nonstructural protein 2; NS1/E2 protein
 C:Species: hepatitis C virus
 A:Variety: isolate JK5
 C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 17-Nov-2000
 C:Accession: S19876
 R:Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
 Submitted to the EMBL Data Library, September 1991
 A:Description: Sequence analysis of putative structural regions of Hepatitis C Virus iso
 A:Reference number: S18029
 A:Accession: S19876
 A:Molecule type: genomic RNA
 A:Residues: 1-782 <HON>
 A:Cross-references: EMBL:X61595; NID:g59486; PIDN:CAA43792.1; PID:g59487
 A:Experimental source: isolate JK5
 C:Superfamily: hepatitis C virus genome polypeptin
 C:Keywords: capsid protein; core protein; envelope protein; glycoprotein; nonstructural
 F:1-191/Product: core protein #status predicted <MAT1>
 F:192-383/Product: envelope protein 1 #status predicted <MAT2>
 F:384-733/Product: NS1/E2 protein #status predicted <MAT3>
 F:734-782/Product: nonstructural protein 2 (fragment) #status predicted <MAT4>

Alignment Scores:

Pred. No.:	1,18e-89	Length:	782
Score:	1063.00	Matches:	196
Percent Similarity:	97.60%	Conservative:	7
Best Local Similarity:	94.23%	Mismatches:	5
Query Match:	92.03%	Indels:	0
DB:	2	Gaps:	0

US-09-899-303A-7 (1-633) x S19876 (1-782)

```

QY 4 TTGGTAGGTCATCGATACCTTACGTGGCGCTTGCACCACTCATGGGATACATTCG 63
    |||
Db 119 LeuGlyValIleAspThrLeuThrCysGlyPheAlaAspProMetGlyTyrIlePro 138
QY 64 CTGTCGGGCCCCCCTAGGGGGTGCCTCCAGAGCCCTGGCGCATGGCTCCGGTCTG 123
    |||
Db 139 LeuValGlyAlaProLeuGlyGlyAlaAlaArgAlaLeuAlaHisGlyValAlaGValLeu 158
QY 124 GAAGAGCGCGCTGAACATGACAGAGGAATTCCTGCTGCTGCTTCCTATCTTCCTC 183
    |||
Db 159 GluAspGlyValAsnIleThrAlaThrGlyAsnLeuProGlySerPheSerIlePheLeu 178
QY 184 TTGGCTTTACTGCTGCTGTGACCAATTCAGCTTCGCTTATGAGTGGCGCAACGTGCC 243
    |||
Db 179 LeuAlaLeuLeuSerCysLeuThrIleProAlaSerAlaTyrGluValAlaGValSer 198
QY 244 GGGATGACCATGTCAGAGAGAGCTGCTCACTCAAGCAATTCGTATGAGGAGCGGAC 303
    |||
Db 199 GlyLeuThrHisValThrAsnAspCysSerAsnSerSerIleValIleGluAlaAlaAsp 218
QY 304 ATGATCATGACACACCCCGGGTGCCTGCGCTTCGCGGAGAAACAATCTCCGCTGC 363
    |||
Db 219 LeuIleMetHisThrProGlyCysValProCysValLeuGluIleAsnSerSerArgCys 238
QY 364 TTGGTAGCGCTCAACCCCGGCTGCGCTGCGCTTCGCGGAGAAACAATCTCCGCTGC 423
    |||
Db 239 TrpValAlaLeuThrProThrLeuAlaAlaArgAsnSerSerIleProThrThrIle 258
QY 424 CGAGCGCAGCTGCATTCCTGCTGCTGCGGCGGCTTCCTGCTTCGCTATGACGTGGG 483
    |||
Db 259 ArgArgHisValAspLeuLeuValIleGlyAlaAlaAlaPheCysSerAlaMetIleValGly 278
QY 484 GATCTGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 543
    |||
Db 279 AspLeuCysGlySerValAlaPheLeuValSerGlnLeuPheThrPheSerProAlaGlyTyr 298
QY 544 GAGACGCTGACAGAGTCGCAATTCCTCAATCTATCCGGCCATCAACAGGTCACCGTATG 603
    |||
Db 299 GluThrValGlnAspCysAsnCysSerLeuTyrSerGlyHisValSerGlyHisArgMet 318
QY 604 GCTTGGATATGATGATGAACCTGG 627
    |||
Db 319 AlaTrpAspMetMetMetAsnTrp 326

```

RESULT 10

S18032
 genome polypeptin - hepatitis C virus (isolate JK4) (fragment)
 N:contains: core protein; envelope protein 1; nonstructural protein 2; NS1/E2 protein
 C:Species: hepatitis C virus
 A:Variety: isolate JK4
 C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 17-Nov-2000
 C:Accession: S18032
 R:Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
 Submitted to the EMBL Data Library, September 1991
 A:Description: Sequence analysis of putative structural regions of Hepatitis C Virus
 A:Reference number: S18029
 A:Accession: S18032
 A:Molecule type: genomic RNA
 A:Residues: 1-782 <HON>
 A:Cross-references: EMBL:X61594
 A:Experimental source: isolate JK4
 C:Superfamily: hepatitis C virus genome polypeptin
 C:Keywords: capsid protein; core protein; envelope protein; glycoprotein; nonstructur
 F:1-191/Product: core protein #status predicted <MAT1>
 F:192-383/Product: envelope protein 1 #status predicted <MAT2>
 F:384-733/Product: NS1/E2 protein #status predicted <MAT3>
 F:734-782/Product: nonstructural protein 2 (fragment) #status predicted <MAT4>

Alignment Scores:

Pred. No.:	1,18e-89	Length:	782
Score:	1063.00	Matches:	194
Percent Similarity:	97.12%	Conservative:	8

Best Local Similarity: 93.27%
Query Match: 92.03%
DB: 2

Mismatches: 6
Indels: 0
Gaps: 0

US-09-899-303a-7 (1-633) x S18032 (1-782)

QY 4 TTGGGTAAAGTCATGATACCTTACGTCGGGCTTCCGCGCATGATGGGTACATTCCG 63
DB 119 LeuGlyValIleAspThrLeuThrCysGlyPheAlaSerLeuMetGlyTyrIlePro 138
QY 64 CTCGTCGGCGCCCGCCAGGGGGTGTGTCGAGACCGTCGGGCGCATGGCGTTCGCTG 123
DB 139 LeuValGlyAlaProIleuGlyGlyAlaAlaArgAlaLeuAlaHisGlyValAlaValLeu 158
QY 124 GAAGACGGCGTGAACATAATGCAACAGGAATTTGCTGCTGCTTCTCTATCTTCCTC 183
DB 159 GluAspGlyAlaAsnTyrAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeu 178
QY 184 TTGGCTTACTGTCTCTGACCATTCAGCTTCGCTTATGAGGTGCGCAACGTGTC 243
DB 179 LeuAlaLeuLeuSerCysLeuThrValProValSerThrTyrGluValAlaArgAsnValSer 198
QY 244 GGGATGACCATGTCACAGAACGACTGTCACATCAAGCATTTGTTATGAGGCGAGCGGAC 303
DB 199 GlyValIleThrValIleThrAsnAspCysSerAsnSerIleValTyrGluAlaAlaAsp 218
QY 304 ATGATCATGCAACACCCCGGGTGGCTGCTGCTGCGAGAACACTCTTCCCGCTGC 363
DB 219 MetIleMetHisThrProGlyCysValProCysValAlaArgGluAsnSerSerArgCys 238
QY 364 TGGGTACCGCTCACCCCGAGCTGCGAGCTAGAACCGCAACGCTTCCCGCTACGACATA 423
DB 239 TrpValAlaLeuThrProThrLeuAlaAlaArgAsnSerSerIleProThrThrIle 258
QY 424 CGAGCGCACGTCGATTTGCTGCTGCTGGGCGGCTTCTTTCGCTATGATGAGTGGG 483
DB 259 ArgArgHisValAlaSerLeuValGlyAlaAlaAlaLeuCysSerAlaMetTyrValGly 278
QY 484 GATCTGTGCGGAGTGTCTCTCTGCTGCTCCAGCTGTTCACCATCTGCGCTCGCGGAT 543
DB 279 AspLeuGlySerValIlePheLeuValSerGlnLeuPheThrPheSerProArgArgTyr 298
QY 544 GAGACGGTGCAGGACTGCATATGCTATCCCGGCGCATACAGGTCACCGATG 603
DB 299 GluThrValGlnAspCysAsnCysSerIleTyrProGlyHisValSerGlyHisArgMet 318
QY 604 GCTTGGATATGATGATGAACCTGG 627
DB 319 AlaTrpAspMetMetMetAsnTrp 326

RESULT 11

S21471
genome polyprotein - hepatitis C virus (fragment)
N:Contains: capsid protein; envelope protein
C:Species: hepatitis C virus
C>Date: 20-Feb-1995 #sequence_revision 19-Apr-1996 #text_change 17-Nov-2000
C:Accession: S21471
R:Mogam, W.K.
submitted to the EMBL Data Library, April 1992
A:Reference number: S21471
A:Accession: S21471
A:Molecule type: genomic RNA
A:Residues: 1-369 <MOG>
A:Cross-references: EMBL:X65924; NID:959466; PIDN:CAA6717.1; PID:959467
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: capsid protein; envelope protein; glycoprotein; polyprotein

Alignment Scores:
Pred. No.: 2.17e-89
Score: 1060.00
Percent Similarity: 97.60%
Best Local Similarity: 94.71%
Query Match: 91.77%

Length: 369
Matches: 197
Conservative: 6
Mismatches: 5
Indels: 0

DB: 2 Gaps: 0

US-09-899-303a-7 (1-633) x S21471 (1-369)

QY 4 TTGGGTAAAGTCATGATACCTTACGTCGGGCTTCCGCGCATGATGGGTACATTCCG 63
DB 119 LeuGlyValIleAspThrLeuThrCysGlyPheAlaSerLeuMetGlyTyrIlePro 138
QY 64 CTCGTCGGCGCCCGCCAGGGGGTGTGTCGAGACCGTCGGGCGCATGGCGTTCGCTG 123
DB 139 LeuValGlyAlaProIleuGlyGlyAlaAlaArgAlaLeuAlaHisGlyValAlaValLeu 158
QY 124 GAAGACGGCGTGAACATAATGCAACAGGAATTTGCTGCTGCTTCTCTATCTTCCTC 183
DB 159 GluAspGlyAlaAsnTyrAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeu 178
QY 184 TTGGCTTACTGTCTCTGACCATTCAGCTTCGCTTATGAGGTGCGCAACGTGTC 243
DB 179 LeuAlaLeuLeuSerCysLeuThrValProValSerThrTyrGluValAlaArgAsnValSer 198
QY 244 GGGATGACCATGTCACAGAACGACTGTCACATCAAGCATTTGTTATGAGGCGAGCGGAC 303
DB 199 GlyValIleThrValIleThrAsnAspCysSerAsnSerIleValTyrGluAlaAlaAsp 218
QY 304 ATGATCATGCAACACCCCGGGTGGCTGCTGCTGCGAGAACACTCTTCCCGCTGC 363
DB 219 ThrIleMetHisThrProGlyCysValProCysValAlaArgGluAsnSerSerArgCys 238
QY 364 TGGGTACCGCTCACCCCGAGCTGCGAGCTAGAACCGCAACGCTTCCCGCTACGACATA 423
DB 239 TrpValAlaLeuThrProThrLeuAlaAlaArgAsnSerSerIleProThrThrIle 258
QY 424 CGAGCGCACGTCGATTTGCTGCTGCTGGGCGGCTTCTTTCGCTATGATGAGTGGG 483
DB 259 ArgArgHisValAlaSerLeuValGlyAlaAlaAlaLeuCysSerAlaMetTyrValGly 278
QY 484 GATCTGTGCGGAGTGTCTCTCTGCTGCTCCAGCTGTTCACCATCTGCGCTCGCGGAT 543
DB 279 AspLeuGlySerValIlePheLeuValSerGlnLeuPheThrPheSerProArgArgTyr 298
QY 544 GAGACGGTGCAGGACTGCATATGCTATCCCGGCGCATACAGGTCACCGATG 603
DB 299 GluThrValGlnAspCysAsnCysSerIleTyrProGlyHisValSerGlyHisArgMet 318
QY 604 GCTTGGATATGATGATGAACCTGG 627
DB 319 AlaTrpAspMetMetMetAsnTrp 326

RESULT 12

GNMVTG
genome polyprotein - hepatitis C virus
N:Contains: capsid protein C; envelope protein M; hepacivirin (PC 3.4.21.98) (nonstru
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 19-Jan-2001
C:Accession: A38465
R:Takamizawa, A.; Mori, C.; Fuke, I.; Manabe, S.; Murakami, S.; Fujita, J.; Onishi, E
J. Virol. 65, 1105-1113, 1991
A:Title: Structure and organization of the hepatitis C virus genome isolated from hum
A:Reference number: A38465; M01D:91140696; PMID:18477440
A:Accession: A38465
A:Molecule type: genomic RNA
A:Residues: 1-3010 <YAK>
A:Cross-references: EMBL:M58335; NID:9329770; PIDN:AA72945.1; PID:9329771
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstruct
F:2-115/Product: capsid protein C #status predicted <CP>
F:116-191/Product: envelope protein M #status predicted <EPK>
F:192-389/Product: major envelope protein E #status predicted <NEE>
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F:1007-1615/Product: hepacivirin #status predicted <NS3>
F:1230-1237/Region: nucleotide-binding motif A (P-loop)

F:1312-1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEXH motif
F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4>
F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4>
F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F:196,209,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,2077,22

Alignment Scores:

Pred. No.:	2,31e-89	Length:	3010
Score:	1060.00	Matches:	194
Percent Similarity:	97.12%	Conservative:	8
Best Local Similarity:	93.27%	Mismatches:	6
Query Match:	91.77%	Indels:	0
DB:	1	Gaps:	0

US-09-899-303a-7 (1-633) x GNMVTC (1-3010)

```
QY 4 TTGGGTAAAGGTCAATGCATACCTTACGTCGGGCTTCGCGACCTCAATGGGTCATTCGCG 63
    |||||
Db 119 LeuGlyValValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIlePro 138
    |||||
QY 64 CTGCTGGCGCCGCCCTAGGGGGTGTGCCAGAGCCCTGGCGATGGCGTCTGCTG 123
    |||||
Db 139 LeuValGlyAlaProLeuGlyGlyAlaAlaArgAlaLeuAlaHisGlyValArgValLeu 158
    |||||
QY 124 GAAGAGGCGCTGAACATGCAACAGGGAATTTGCTGGTCTCTCTTCTATCTTCCTC 183
    |||||
Db 159 GluAspGlyAlaAsnTyrAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeu 178
    |||||
QY 184 TTGGCTTTACTGTCTGCTGTGACCAATTCACGCTTCGCTATAGAGGCGCCACGTCCTC 243
    |||||
Db 179 LeuAlaLeuLeuSerCysLeuThrThrProAlaSerAlaTyrGluValHisAsnValSer 198
    |||||
QY 244 GGGATGATACCATGTCACAGAGCACTGCTCCAACTCAAGCATTTGTATGAGGACGCGAC 303
    |||||
Db 199 GlyIleThrHisValIleThrAsnAspCysSerAsnAlaSerIleValTyrGluAlaAlaAsp 218
    |||||
QY 304 ATGATCATGACACACCCCGGGTGGCTGCTGCTGCTGGGAGAAACAATCTTCCCGTGC 363
    |||||
Db 219 LeuIleMetHisThrProGlyCysValProCysValArgGluValHisSerSerIleProCys 238
    |||||
QY 364 TTGGGTACCGCTCAACCCCAAGCTGCGAGCTGAGAGCAAGCCGACGCTCCCACTAGACATA 423
    |||||
Db 239 TrpValAlaLeuThrProThrLeuAlaAlaArgAsnValThrIleProThrThrIle 258
    |||||
QY 424 CGACGCCACGTCGATTTGCTGTTGGGGCGGCTTCTGTTCCGCTATGACGTGGGG 483
    |||||
Db 259 ArgArgHisValAspLeuLeuValGlyAlaAlaAlaPheCysSerAlaMetTyrValGly 278
    |||||
QY 484 GATCTGCGGATGCTGCTCTGCTGCTGCTCCACCTGTCACCAATTCGCGCTCCCGGCAAT 543
    |||||
Db 279 AspLeuCysGlySerValPheLeuValSerGlnLeuPheThrPheSerProAlaArgHis 298
    |||||
QY 544 GAGACGGTGCAGAGCTGCAATTTGCTCAATCTATCCGGCCACATTAACAGGTCAACGGTATG 603
    |||||
Db 299 ValThrLeuGlnAspCysAsnCysSerIleTyrProGlyHisValSerGlyHisArgMet 318
    |||||
QY 604 GCTTGGGATATGATGATGAACCTGG 627
    |||||
Db 319 AlaTrpAspMetMetMetAsnTrp 326
    |||||
```

RESULT 13

JQ1925
polyprotein - hepatitis C virus (isolate HCV-KE)
N:contains: C protein; E1 protein; E2/NS1 protein
C:Species: hepatitis C virus
C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 17-Nov-2000
C:Accession: JQ1925
R:Abbe, K.; Inchauspe, G.; Fujisawa, K.
J. Gen. Virol. 73, 2725-2729, 1992
A:Title: Genomic characterization and mutation rate of hepatitis C virus isolated from a
A:Reference number: JQ1925; MUID:93019030; PMID:1383400
A:Accession: JQ1925

A:molecule type: mRNA
A:Residues: 1-520 <ABE>
A:Cross-references: DDBJ:D10687; NID:9221544; PIDN:BA01529.1; PID:9221545
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: polyprotein; transmembrane protein
F:1-191/Product: C protein #status predicted <CPR>
F:192-383/Product: E1 protein #status predicted <E1R>
F:384-520/Product: E2/NS1 protein #status predicted <E2P>

Alignment Scores:

Pred. No.:	4,15e-89	Length:	520
Score:	1057.00	Matches:	195
Percent Similarity:	96.63%	Conservative:	6
Best Local Similarity:	93.75%	Mismatches:	7
Query Match:	91.52%	Indels:	0
DB:	2	Gaps:	0

US-09-899-303a-7 (1-633) x JQ1925 (1-520)

```
QY 4 TTGGGTAAAGGTCAATGCATACCTTACGTCGGGCTTCGCGACCTCAATGGGTCATTCGCG 63
    |||||
Db 119 LeuGlyValValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIlePro 138
    |||||
QY 64 CTGCTGGCGCCGCCCTAGGGGGTGTGCCAGAGCCCTGGCGATGGCGTCTGCTG 123
    |||||
Db 139 LeuValGlyAlaProLeuGlyGlyAlaAlaArgAlaLeuAlaHisGlyValArgValLeu 158
    |||||
QY 124 GAAGAGGCGCTGAACATGCAACAGGGAATTTGCTGGTCTCTCTTCTATCTTCCTC 183
    |||||
Db 159 GluAspGlyAlaAsnTyrAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeu 178
    |||||
QY 184 TTGGCTTTACTGTCTGCTGTGACCAATTCACGCTTCGCTATAGAGGCGCCACGTCCTC 243
    |||||
Db 179 SerAlaLeuMetSerCysLeuThrThrProAlaSerAlaTyrGluValArgAsnValSer 198
    |||||
QY 244 GGGATGATACCATGTCACAGAGCACTGCTCCAACTCAAGCATTTGTATGAGGACGCGAC 303
    |||||
Db 199 GlyIleThrHisValIleThrAsnAspCysSerAsnAlaSerIleValTyrGluAlaAlaGly 218
    |||||
QY 304 ATGATCATGACACACCCCGGGTGGCTGCTGCTGCTGGGAGAAACAATCTTCCCGTGC 363
    |||||
Db 219 MetIleMetHisThrProGlyCysValProCysValArgGluValHisSerSerIleProCys 238
    |||||
QY 364 TTGGGTACCGCTCAACCCCAAGCTGCGAGCTGAGAGCAAGCCGACGCTCCCACTAGACATA 423
    |||||
Db 239 TrpValAlaLeuThrProThrLeuAlaAlaArgAsnAlaSerValProThrThrIle 258
    |||||
QY 424 CGACGCCACGTCGATTTGCTGTTGGGGCGGCTTCTGTTCCGCTATGACGTGGGG 483
    |||||
Db 259 ArgArgHisValAspLeuLeuValGlyAlaAlaAlaThrLeuCysSerAlaMetTyrValGly 278
    |||||
QY 484 GATCTGCGGATGCTGCTCTGCTGCTGCTCCACCTGTCACCAATTCGCGCTCCCGGCAAT 543
    |||||
Db 279 AspLeuCysGlySerValPheLeuValSerGlnLeuPheThrPheSerProAlaArgHis 298
    |||||
QY 544 GAGACGGTGCAGAGCTGCAATTTGCTCAATCTATCCGGCCACATTAACAGGTCAACGGTATG 603
    |||||
Db 299 GluThrValGlnAspCysAsnCysSerIleTyrProGlyHisValSerGlyHisArgMet 318
    |||||
QY 604 GCTTGGGATATGATGATGAACCTGG 627
    |||||
Db 319 AlaTrpAspMetMetMetAsnTrp 326
    |||||
```

RESULT 14

JH0711
genome polyprotein - hepatitis C virus (strain PKC1) (fragments)
N:contains: envelope protein E1; envelope protein E2; nonstructural protein NS1; nons
C:Species: hepatitis C virus
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 17-Nov-2000
C:Accession: JH0711
R:Lin, K.; Hu, Z.; Li, H.; Prince, A.M.; Inchauspe, G.
Gene 114, 245-250, 1992
A:Title: Genomic typing of hepatitis C viruses present in China.

A:Reference number: JH0711; MUID:92290283; PMID:1318245
A:Accession: JH0711
A:Molecule type: genomic RNA
A:Residues: 1-550 <LID>
A:Cross-references: GB:M74888; GB:M74889
A:Note: the nucleotide sequence is not complete
A:Note: translation of the nucleotide sequence is not complete
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: envelope protein; glycoprotein; nonstructural protein; nucleocapsid; polyp
F:1-190/Product: nucleocapsid protein C #status predicted <CP>
F:191-380/Product: envelope protein E1 #status predicted <EP>
F:381-514/Product: envelope protein E2 or nonstructural protein NS1 (fragment) #status F
F:515-550/Product: nonstructural protein NS5 (fragment) #status predicted <NS>
F:196,233,250,305,416,422,429,447/Binding site: carbohydrate (Asn) (covalent) #status pr

Alignment Scores:		
Pred. No.:	5,28e-88	Length: 550
Score:	1045.00	Matches: 188
Percent Similarity:	95.19%	Conservative: 10
Best Local Similarity:	90.38%	Mismatches: 10
Query Match:	90.48%	Indels: 0
DB:	2	Gaps: 0

JS-09-899-303A-7 (1-633) x JH0711 (1-550)

OY	4	TTGGGTAAAGGCATCATGATACACCTTACGTCGGGCTTCGGCGACCTCATATGGGTAACATTTCGG	63
Db	119	LeuGIyIvsValIIAspPrIthLeuthrCysGIyPheIIAspLeuMeGIyIytrIIePro	138
OY	64	CTCGTCGGCGCCCTTATAGGGGGTGCCTCCAGAGCGCCCTGGGCGCATAGGCGCGGTTGTG	123
Db	139	LeuValIGIyIaIaProIeuGIyGIyValValaIaArgIaIeuaIIaHIsGIyValaIyGIaIIeu	158
OY	124	GAAGACGGCGGTGACTATATGCACAGGGAATTGGCTGGTGTCTTCTATCTTCTCTC	183
Db	159	GIuAspGIyValaIIAspIyIaIaIhCGLyAsnLeuProGIyCysSerPheSerIIePheIeu	178
OY	184	TTGGCTTACATGCTCTGCTGTGCACATTCAGCTTCAGCTTCGCGCTTATAGAGTGCAGACCTGTC	243
Db	179	LeuIIaIaIeLeuSerCysLeuIthrIthProIIaSerIaIaIyGIyValaIaArgAsnIIaSer	198
OY	244	GAGATGTACCATGTGCAGACAGACTGTCCAACTCAGACCATTTGTATAGGCGACGGAC	303
Db	199	GIyValIyIhIIAspIyIaIhAsnAspCysSerAsnIthAsnIIeValIyGIyIthrIIaIAsp	218
OY	304	ATGATCATGCACACCCCGGGGCGCTGCGCTTCGCTGGGAGAACAACTCTTCCGGCTGC	363
Db	219	MeIIeMeIIhIsthrProGIyCysIaIeProCysValaIaArgAspAsGIyIthSerIArgCys	238
OY	364	TGGGAGGCGGCTACCCCGACGCTGCAGCTAGAGAACGCGAGGCGTCCCACTACGACATTA	423
Db	239	TrpValaIaIeIthrProIthIeuaIIaIaArgAsnIIaSerIIeProIthIhIaIaIle	258
OY	424	CGAGCCAGCTGCATTTCTCTGTTGGGGCGGCTGCTTCTGTCTGCTAGTACGTGGGG	483
Db	259	ArgArgIIhIsthrAspLeuIeuaIaIyIaIaIaIthrPheCysSerIIaIeMeIyValGIy	278
OY	484	GATCTCTCGGATCTGTCCTCTGCTCCGACAGCTTCACACATTCGCGCTCCCGGAT	543
Db	279	AspLeuCysGIySerIIePheIeuaIaIserGIuIeuaIePheIthrPheSerProArgGIhIs	298
OY	544	GAGAGGTGCGAGATGCAATTTGTCATTTATCCCGGCGACATTAACAGGTACACCGTATG	603
Db	299	GIuIthrValGIuAspCysAsnCysSerIIeIytrProGIyIIhIsValIthGIyIIhIsArgMet	318
OY	604	GCTTGGATATGATGATGAATCGG	627
Db	319	AlaItrAspMeIIeMeIAsnItr	326
RESULT	15		
	301926		
polyprotein - hepatitis C virus (isolate HCV-476)			
N:contains: C protein; E1 protein; E2/NS1 protein			

```

C:Species: hepatitis C virus
C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 17-Nov-2000
C:Accession: J01926
R:Abel, K.; Inchausti, G.; Fujisawa, K.
J. Gen. Virol. 73, 2725-2729, 1992
A:Title: Genomic characterisation and mutation rate of hepatitis C virus isolated from
A:Reference number: J01925; MUID:93019030; PMID:1383400
A:Accession: J01926
A:Molecule type: mRNA
A:Residues: 1-523 <ABE>
A:Cross-references: DDBJ:J01687
C:Superfamily: hepatitis C virus genome polypeptide
C:Keywords: polypeptide
F:1192-383/Product: C protein #status predicted <CPR>
F:192-383/Product: E1 protein #status predicted <E1R>
F:384-523/Product: E2/NS1 protein #status predicted <E2P>

Alignment Scores:
Pred. No.:      6,52e-88      Length:      523
Score:          1044.00      Matches:     192
Percent Similarity: 96.15%      Conservative: 8
Best Local Similarity: 92.31%      Mismatches: 8
Query Match:    90.39%      Indels:     0
DB:              2          Gaps:       0

```

US-09-899-303A-7 (1-63) x J01926 (1-523)

QY 4 TTGGGTAAAGTCATACGATATACCCCTTACGTGGCGGCTTCGCGCACTCATAGGGGTACATTCG 63

DB 119 LeuGlyIysValIleAspThrLeuThrCysGlyPheIleAspLeuMetGlyTyrIlePro 138

QY 64 CTCGTCGGCGGCCCCCTAGGGGGGTGCTGCCAGAGCCCTGGGGGCATAGGGGTCCGGGTTCTG 123

DB 139 LeuValIleGlyAlaProLeuGlyIleValAspThrAlaLeuIleAlaIleGlyValAlaGlyLeu 158

QY 124 GAGACGGCGGTCAACTATGACAACAGAGAAATTGACCTGGCTGCTCTTCTATATGTTCTC 183

DB 159 GluAspGlyValAsnTyrAlaThrIleGlyAsnLeuProGlyCysSerPheSerIlePheLeu 178

QY 184 TTGGCTTACGTCTGCTGTCTGACCAATCCAGCTCCGCTTATGAGTGGCGCAAGCTGTCC 243

DB 179 SerAlaLeuMetSerCysLeuThrAlaProIleSerAlaTyrGlyValAlaTyrAsnValSer 198

QY 244 GGGATGTACCAATGTACAGACAGACATGTCTCAACTCAACCAATTTGTATGAGGACGGCAC 303

DB 199 GlyIleTyrPheIleValThrAsnProCysSerAsnSerIleAlaTyrGlyAlaIleGly 218

QY 304 ATGATCATGACACACCCCGGGTGCCTGCTGCCCTGCTGGGAGAACACTTCTCCGGCTGC 363

DB 219 MetIleMetHisThrProGlyValProCysValAlaArgIleAspAsnSerSerAlaGly 238

QY 364 TGGGTAGCGGCTCACCCCCAGCGCTGCACTAGGAACGCCAGCGTCCCACTACGACAATA 423

DB 239 TrpValAlaLeuThrProThrLeuAlaIleAlaTyrAsnAlaSerValProThrThrThrIle 258

QY 424 CGACCCCAAGTCGATTTCTGCTGGGGCGGGCTGTTCTGTCCGCTATGACGTGGGG 483

DB 259 ArgAlaGlnHisValAspLeuLeuValGlyAlaAlaThrIleCysSerAlaMetGlyValGly 278

QY 484 GATCTCTCGGATCTGTCTCTCCCTGCTCCAGCTGTACCAACTCTGCGCGCGCAT 543

DB 279 AspLeuCysGlySerValPheLeuValSerGlnLeuIleThrPheSerProAlaArgTyr 298

QY 544 GAGAGCGGTGCAGAGATGCAATTGCTCATCTATCCCGGSCACATACAGAGTACCGATG 603

DB 299 GluThrValGlnAspCysAsnProCysSerLeuTyrProGlyIleValSerGlyIleAspMet 318

QY 604 GCTTGCGATATGATGATGAATG 627

DB 319 AlaTrpAspMetMetMetAsnTrp 326

Search completed: February 12, 2003, 08:57:19

Job time : 36.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 12, 2003, 08:55:54 ; Search time 241 Seconds
(without alignments)
3386.855 Million cell updates/sec

Title: US-09-899-303A-7
Perfect score: 1155
Sequence: 1 ATGTTGGTGAAGTCATCGA.....TGATGATGAACGTGTAATAG 633

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 9138288

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODE=frame_n2p.model -DEV=xlp
-O/cgn2_1/USPTO.spool/US09899303.r/unat.12022003.085450.28326/app.query.fasta.1.775
-DB=Pending_Patents_AA_Main -QPM=fastan -SUFFIX=n2p.rapm -MINMATCH=0.1
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blousum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pcpt -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09899303.@CGN.1.1.175.@unat.12022003.085450.28326
-NCPU=6 -ICPU=3 -NO_XLPHY -NO_MMAP -LARGEIOBURY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -MARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -IGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/1/paa/PCTUS.COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06.COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07.COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08.COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US081.COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US082.COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US083.COMB.pep:*
8: /cgn2_6/ptodata/1/paa/US084.COMB.pep:*
9: /cgn2_6/ptodata/1/paa/US085.COMB.pep:*
10: /cgn2_6/ptodata/1/paa/US086.COMB.pep:*
11: /cgn2_6/ptodata/1/paa/US087.COMB.pep:*
12: /cgn2_6/ptodata/1/paa/US088.COMB.pep:*
13: /cgn2_6/ptodata/1/paa/US089.COMB.pep:*
14: /cgn2_6/ptodata/1/paa/US090.COMB.pep:*
15: /cgn2_6/ptodata/1/paa/US091.COMB.pep:*
16: /cgn2_6/ptodata/1/paa/US092.COMB.pep:*
17: /cgn2_6/ptodata/1/paa/US093.COMB.pep:*
18: /cgn2_6/ptodata/1/paa/US094.COMB.pep:*
19: /cgn2_6/ptodata/1/paa/US095.COMB.pep:*
20: /cgn2_6/ptodata/1/paa/US096.COMB.pep:*
21: /cgn2_6/ptodata/1/paa/US097.COMB.pep:*
22: /cgn2_6/ptodata/1/paa/US098.COMB.pep:*
23: /cgn2_6/ptodata/1/paa/US099.COMB.pep:*
24: /cgn2_6/ptodata/1/paa/US100.COMB.pep:*
25: /cgn2_6/ptodata/1/paa/US101.COMB.pep:*
26: /cgn2_6/ptodata/1/paa/US102.COMB.pep:*
27: /cgn2_6/ptodata/1/paa/US60.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1118	96.8	209	22	US-09-899-303-8	Sequence 8, Appl1
2	1118	96.8	209	22	US-09-899-303A-8	Sequence 8, Appl1
3	1118	96.8	209	23	US-09-973-025-8	Sequence 8, Appl1
4	1118	96.8	209	23	US-09-995-791-8	Sequence 8, Appl1
5	1118	96.8	209	23	US-09-995-808-8	Sequence 8, Appl1
6	1118	96.8	209	23	US-09-995-860-8	Sequence 8, Appl1
7	1118	96.8	209	24	US-10-020-510-8	Sequence 8, Appl1
8	1118	96.8	209	25	US-10-128-590-88	Sequence 8, Appl1
9	1113	96.4	263	22	US-09-899-303-6	Sequence 6, Appl1
10	1113	96.4	263	22	US-09-899-303A-6	Sequence 6, Appl1
11	1113	96.4	263	23	US-09-973-025-6	Sequence 6, Appl1
12	1113	96.4	263	23	US-09-995-791-6	Sequence 6, Appl1
13	1113	96.4	263	23	US-09-995-808-6	Sequence 6, Appl1
14	1113	96.4	263	23	US-09-995-860-6	Sequence 6, Appl1
15	1113	96.4	263	24	US-10-020-510-6	Sequence 6, Appl1
16	1109	96.0	210	22	US-09-899-303-14	Sequence 6, Appl1
17	1109	96.0	210	22	US-09-899-303A-14	Sequence 14, Appl1
18	1109	96.0	210	23	US-09-973-025-14	Sequence 14, Appl1
19	1109	96.0	210	23	US-09-995-791-14	Sequence 14, Appl1
20	1109	96.0	210	23	US-09-995-808-14	Sequence 14, Appl1
21	1109	96.0	210	23	US-09-995-860-14	Sequence 14, Appl1
22	1109	96.0	210	24	US-10-020-510-14	Sequence 14, Appl1
23	1108	95.9	692	22	US-09-899-303A-48	Sequence 48, Appl1
24	1108	95.9	692	22	US-09-899-303-48	Sequence 48, Appl1
25	1108	95.9	692	23	US-09-973-025-48	Sequence 48, Appl1
26	1108	95.9	692	23	US-09-995-791-48	Sequence 48, Appl1
27	1108	95.9	692	23	US-09-995-808-48	Sequence 48, Appl1
28	1108	95.9	692	23	US-09-995-860-48	Sequence 48, Appl1
29	1108	95.9	692	24	US-10-020-510-48	Sequence 48, Appl1
30	1108	95.9	809	22	US-09-899-303-50	Sequence 50, Appl1
31	1108	95.9	809	22	US-09-899-303A-50	Sequence 50, Appl1
32	1108	95.9	809	23	US-09-973-025-50	Sequence 50, Appl1
33	1108	95.9	809	23	US-09-995-791-50	Sequence 50, Appl1
34	1108	95.9	809	23	US-09-995-808-50	Sequence 50, Appl1
35	1108	95.9	809	23	US-09-995-860-50	Sequence 50, Appl1
36	1108	95.9	809	24	US-10-020-510-50	Sequence 50, Appl1
37	1094	94.7	450	10	US-08-635-886B-188	Sequence 188, App
38	1094	94.7	450	10	US-08-635-886B-188	Sequence 188, App
39	1094	94.7	513	3	US-07-790-382A-31	Sequence 31, Appl1
40	1094	94.7	733	3	US-07-866-045-12	Sequence 12, Appl1
41	1092	94.5	3010	19	US-09-539-601-3	Sequence 3, Appl1
42	1092	94.5	3010	19	US-09-539-601-21	Sequence 21, Appl1
43	1092	94.5	3010	19	US-09-539-601-27	Sequence 27, Appl1
44	1092	94.5	3010	19	US-09-539-601-33	Sequence 33, Appl1
45	1081	93.6	3010	20	US-09-662-454-3	Sequence 3, Appl1

ALIGNMENTS

RESULT 1
US-09-899-303-8
Sequence 8, Application US/09899303
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
BOSMAN, FONS
DE MARTYNOFF, GUY
BUYSE, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYTE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON

STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIDM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/899,303
FILING DATE: 06-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/612,973
FILING DATE: 11-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4100
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 209 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-899-303-8
Alignment Scores:
Pred. No.: 5,77e-96 Length: 209
Score: 1118.00 Matches: 209
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 96.80% Indels: 0
DB: 22 Gaps: 0
US-09-899-303A-7 (1-633) x US-09-899-303-8 (1-209)
QY 1 ATGTTGGGTAAAGGTCATGATACCTTACGTGGCGCTTCCGAGCTTCACCAATTCGCGCCGG 60
DB 1 MetLeuGlyLysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIle 20
QY 61 CCGCTGTCGGGCGCCCGCTAGGGGGTGTGCCAGAGCCCTGGCGCATGGCGCGGT 120
DB 21 ProLeuValIGlyAlaProLeuGlyGlyAlaAlaArgAlaLeuAlaHisGlyValArgVal 40
QY 121 CTGGAGAGGGGGGTGAATGCAACAGAGGAATTTGCTGGTTCCTTTCTATCTTC 180
DB 41 LeuGlnAspGlyValAsnThrGlyAsnLeuProGlyCysSerPheSerIlePhe 60
QY 181 CTTTGGCTTTACTGTCGTGTGACCAATTCAGCTTCCGCTTATGAGGCGCAACGTG 240
DB 61 LeuLeuAlaLeuLeuSerLysLeuThrIleProAlaSerAlaTyrGluValArgAsnVal 80
QY 241 TCCGGGATGTACCATGTCAAGAGCACTGCTCCAACTCAAGCATTTGTATGAGGAGCG 300
DB 81 SerGlyMetTyrHisValIleHisAsnAspCysSerAsnSerSerIleValIleGluAlaIle 100
QY 301 GACATGATATGACACACCCCGGGTCCGCGCTTCCGCGGAGACAACTCTTCCCGC 360
DB 101 AspMetIleMetHisThrProGlyCysValProCysValArgGlnAsnAsnSerSerArg 120
QY 361 TCGTGGGTAGCGTCAACCCCAAGCTCCAGCTAGGAGACCGACGAGCTCCCACTAGACA 420
DB 121 CysTrpValAlaLeuThrThrThrLeuAlaAlaArgAsnAlaSerValProThrThrThr 140
QY 421 ATAGACAGCCACGATGATGCTGCTGGGGCGGCTTCTGTTCGCTATGTAAGCTG 480
DB 141 IleArgArgHisValAspLeuLeuValIGlyAlaAlaAlaPheCysSerAlaMetGlyVal 160

QY 481 GGGGATCTCGCGGATCTGTCCTTCGTCGTCCTCCAGCTGTTCACCAATTCGCGCCGG 540
DB 161 GlyAspLeuCysGlySerValPheLeuValSerGlnLeuPheThrIleSerProArgArg 180
QY 541 CATGAGCGGTGGAGCACTGATGCTCAATCTATCCGCGGACCAATACAGTCAACCG 600
DB 181 HisGlnThrValIleAsnAspCysAsnGlySerIleTyrProGlyHisIleThrGlyHisArg 200
QY 601 ATGGCTTGGGATATGATGATGATGATG 627
DB 201 MetAlaTrpAspMetMetMetLanTrp 209
RESULT 2
US-09-899-303A-8
Sequence 8, Application US/09899303A
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
BOSMAN, PONS
DE MARTYNOFF, GUY
BOUYSE, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIDM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/899,303A
FILING DATE: 06-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/612,973
FILING DATE: 11-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4100
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 209 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-899-303A-8
Alignment Scores:
Pred. No.: 5,77e-96 Length: 209
Score: 1118.00 Matches: 209
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 96.80% Indels: 0
DB: 22 Gaps: 0
US-09-899-303A-7 (1-633) x US-09-899-303A-8 (1-209)
QY 1 ATGTTGGGTAAAGGTCATGATACCTTACGTGGCGCTTCCGAGCTTCACCAATTCGCGCCGG 60
DB 1 MetLeuGlyLysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIle 20
QY 61 CCGCTGTCGGGCGCCCGCTAGGGGGTGTGCCAGAGCCCTGGCGCATGGCGCGGT 120


```
FILE REFERENCE: 2551-68
CURRENT APPLICATION NUMBER: US/09/995,791
CURRENT FILING DATE: 2001-11-29
NUMBER OF SEQ ID NOS: 122
SOFTWARE: PatentIn 3.1
SEQ ID NO 8
LENGTH: 209
TYPE: PRF
ORGANISM: Hepatitis C virus
US-09-995-791-8

Alignment Scores:
Pred. No.: 5 77e-96 Length: 209
Score: 1118.00 Matches: 209
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 96.80% Indels: 0
DB: 23 Gaps: 0

US-09-899-303A-7 (1-633) x US-09-995-791-8 (1-209)

QY 1 ATGTGGGTAAGGTCATCGATACCTTACGTGGGCTTCCGACCTCATGGGGTACATT 60
DB 1 MetLeuGlyLysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIle 20
QY 61 CCGCTGTCGGGCCCCCCTAGGGGGTGTGCCAGAGCCCTGGCGCATGGCGGT 120
DB 21 ProLeuValGlyAlaProLeuGlyGlyAlaAlaArgAlaLeuAlaHisGlyValArgVal 40
QY 121 CTGGAAGAGCGGCGTAACATGCAACAGGGAATTTGCTGCTTCTTCTATCTTC 180
DB 41 LeuGluAspGlyValAsnTyrAlaThrGlyAsnLeuProGlyCysSerPheSerIlePhe 60
QY 181 CTCCTGGCTTACTGCTGCTGTGACCAATTCAGCTTCGCTTATGAGGCGACGG 240
DB 61 LeuLeuAlaLeuLeuSerCysLeuThrIleProAlaSerAlaTyrGluValArgAsnVal 80
QY 241 TCCGGGATGTACCATGTCAAGAACGACTGCTCCAACTCAAGCATTTGTATGAGCGCG 300
DB 81 SerGlyMetTyrHisValIleThrAsnAspCysSerAsnSerSerIleValTyrGluAlaIle 100
QY 301 GACATGATCATGCACACCCCGGGGTGCGCCCTGCGTTCCGGAGAACAACTTCTCCCG 360
DB 101 AspMetIleMetHisThrProGlyCysValProCysValArgGluAsnAsnSerSerArg 120
QY 361 TGTGGGTAGCGCTCACCCCAAGCTCGAGCTAGGAACGCCAGCTCCCACTACGACA 420
DB 121 CysTrpValAlaLeuThrProThrIleuAlaAlaArgAlaAlaSerValProThrThrThr 140
QY 421 ATACGACGCCAGCTGCATTTGCTCGTTGGGGGCGCTGCTTTCTGTTCCGCTATGACGTG 480
DB 141 IleArgArgHisValAspLeuLeuValGlyAlaAlaAlaPheCysSerAlaMetTyrVal 160
QY 481 GGGGATCTGGGGATCTGCTTCTCCGCTCCAGCTGTTCACCACTTCGCGCTGCCGG 540
DB 161 GlyAspLeuCysGlySerValPheLeuValSerGlnLeuPheThrIleSerProArgArg 180
QY 541 CATGAGACGGTCCAGACTGCAATTGCTCAATCTATCCGGCCACATAACAGTCAACGT 600
DB 181 HisGluThrValGlnAspCysAsnSerIleTyrProGlyHisIleThrGlyHisArg 200
QY 601 ATGGCTTGGGATATGATGATGAACCTGG 627
DB 201 MetaIatPrAspMetMetMetAsnTrp 209

RESULT 5
US-09-995-808-8
; Sequence 8, Application US/09995808
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: Purified hepatitis C virus envelope proteins for diagnostic and
; TITLE OF INVENTION: therapeutic use.
; FILE REFERENCE: 2551-70
```

```
CURRENT APPLICATION NUMBER: US/09/995,808
CURRENT FILING DATE: 2001-11-29
NUMBER OF SEQ ID NOS: 122
SOFTWARE: PatentIn 3.1
SEQ ID NO 8
LENGTH: 209
TYPE: PRF
ORGANISM: Hepatitis C virus
US-09-995-808-8

Alignment Scores:
Pred. No.: 5 77e-96 Length: 209
Score: 1118.00 Matches: 209
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 96.80% Indels: 0
DB: 23 Gaps: 0

US-09-899-303A-7 (1-633) x US-09-995-808-8 (1-209)

QY 1 ATGTGGGTAAGGTCATCGATACCTTACGTGGGCTTCCGACCTCATGGGGTACATT 60
DB 1 MetLeuGlyLysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIle 20
QY 61 CCGCTGTCGGGCCCCCCTAGGGGGTGTGCCAGAGCCCTGGCGCATGGCGGT 120
DB 21 ProLeuValGlyAlaProLeuGlyGlyAlaAlaArgAlaLeuAlaHisGlyValArgVal 40
QY 121 CTGGAAGAGCGGCGTAACATGCAACAGGGAATTTGCTGCTTCTTCTATCTTC 180
DB 41 LeuGluAspGlyValAsnTyrAlaThrGlyAsnLeuProGlyCysSerPheSerIlePhe 60
QY 181 CTCCTGGCTTACTGCTGCTGTGACCAATTCAGCTTCGCTTATGAGGCGACGG 240
DB 61 LeuLeuAlaLeuLeuSerCysLeuThrIleProAlaSerAlaTyrGluValArgAsnVal 80
QY 241 TCCGGGATGTACCATGTCAAGAACGACTGCTCCAACTCAAGCATTTGTATGAGCGCG 300
DB 81 SerGlyMetTyrHisValIleThrAsnAspCysSerAsnSerSerIleValTyrGluAlaIle 100
QY 301 GACATGATCATGCACACCCCGGGGTGCGCCCTGCGTTCCGGAGAACAACTTCTCCCG 360
DB 101 AspMetIleMetHisThrProGlyCysValProCysValArgGluAsnAsnSerSerArg 120
QY 361 TGTGGGTAGCGCTCACCCCAAGCTCGAGCTAGGAACGCCAGCTCCCACTACGACA 420
DB 121 CysTrpValAlaLeuThrProThrIleuAlaAlaArgAlaAlaSerValProThrThrThr 140
QY 421 ATACGACGCCAGCTGCATTTGCTCGTTGGGGGCGCTGCTTTCTGTTCCGCTATGACGTG 480
DB 141 IleArgArgHisValAspLeuLeuValGlyAlaAlaAlaPheCysSerAlaMetTyrVal 160
QY 481 GGGGATCTGGGGATCTGCTTCTCCGCTCCAGCTGTTCACCACTTCGCGCTGCCGG 540
DB 161 GlyAspLeuCysGlySerValPheLeuValSerGlnLeuPheThrIleSerProArgArg 180
QY 541 CATGAGACGGTCCAGACTGCAATTGCTCAATCTATCCGGCCACATAACAGTCAACGT 600
DB 181 HisGluThrValGlnAspCysAsnSerIleTyrProGlyHisIleThrGlyHisArg 200
QY 601 ATGGCTTGGGATATGATGATGAACCTGG 627
DB 201 MetaIatPrAspMetMetMetAsnTrp 209

RESULT 6
US-09-995-860-8
; Sequence 8, Application US/09995860
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: Purified hepatitis C virus envelope proteins for diagnostic an
; TITLE OF INVENTION: therapeutic use.
; FILE REFERENCE: 2551-69
; CURRENT APPLICATION NUMBER: US/09/995,860
```

```

: CURRENT FILING DATE: 2001-11-29
: NUMBER OF SEQ ID NOS: 122
: SOFTWARE: PatentIn 3.1
: SEQ ID NO 8
: LENGTH: 209
: TYPE: PRT
: ORGANISM: Hepatitis C virus
US-09-995-860-8

```

Alignment Scores:	
Pred. No.:	5,77e-96
Score:	1118.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	96.80%
DB:	23
Length:	209
Matches:	209
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-899-303A-7 (1-633) x US-09-995-860-8 (1-209)

QY	1	ATGTGGGTAGAGGCATCATGATACCCCTTAAGCGGGCTTGGCCGACCTCATGGGGTACATT	60
Db	1	MetLeuGlyValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIle	20
QY	61	CCGCTCGTGGCGCCCTTAGGGGGGTCTCCAGAGCCCTGGCGCATGGCGTCCGGGT	120
Db	21	ProLeuValGlyAlaProLeuGlyValAlaAlaArgAlaLeuAlaHisGlyValArgVal	40
QY	121	CTGGAACAGCGCCGTAACATAGCAACAAGGAATTGGCTGGTTCCTTTCTCATCTTC	180
Db	41	LeuIuAspGlyValAlaAsnTyrAlaThrGlyAsnLeuProGlyCysSerPheSerIlePhe	60
QY	181	CTCTTGGCTTACCTAGTCCCTGTCTGACCAATTCACGCTTCGCGCTTATGAGAGCGCACGTG	240
Db	61	LeuLeuAlaLeuLeuSerCysLeuThrIleLeuAlaSerAlaTyrGlyValArgAsnVal	80
QY	241	TCCGGGATGATACCATGTGCAGAACAGACTCTCCAACTGAACATTTGTATAGGACGCG	300
Db	81	SerGlyMetTyrHisValThrAsnAspCysSerAsnSerIleValTyrGlyAlaAla	100
QY	301	GACATGATCATGACACACCCCGGGTGGCGCCGCTTCGGGGAGAACACTGTCCCGC	360
Db	101	AspMetIleMetHisThrProGlyCysValProCysValArgGlyAsnAsnSerSerArg	120
QY	361	TGCTGGGTAGGCGTCACACCCGACGCTGAGCTAGGAAGCGCAGCGTCCCACTAGCACA	420
Db	121	CysThrValAlaLeuThrProThrLeuAlaAlaArgAsnAlaSerValProThrThrThr	140
QY	421	ATACGAGCCACGTCGATTTGCTGTTGGGGCGGCTCTTTCTGTCCGSTATGTACGTG	480
Db	141	IleArgArgHisValAspLeuLeuValGlyAlaAlaAlaPheCysSerAlaMetTyrVal	160
QY	481	GGGGATCTCTCGGATGTCGTTCTCTCGGTCCGAGGTTCACCAATCTGGCCGCGCGG	540
Db	161	GlyAspLeuCysGlySerValPheLeuValSerIleLeuPheThrIleSerProAlaArg	180
QY	541	CATGAGAGGTCGAGGACTGCACATTTGCTCAATCTATACCGGCGCCACATACAGGTCACCGT	600
Db	181	HisGlyThrValGlnAspCysAsnGlySerIleTyrProGlyHisIleThrGlyHisArg	200
QY	601	ATGGCTTGGGATATGATGATGAATCGG 627	
Db	201	MetAlaTrpAspMetMetLeuAsnTrrp 209	

RESULT 7
US-10-020-510-8

```

: Sequence 8, Application US/10020510
:
: GENERAL INFORMATION:
:
: APPLICANT: Innogenetics N.V.
:
: TITLE OF INVENTION: Purified hepatitis C virus envelope proteins for diagnostic and
:
: TITLE OF INVENTION: therapeutic use.
:
: FILE REFERENCE: 2551-72
:
: CURRENT APPLICATION NUMBER: US/10/020,510
:
: CURRENT FILING DATE: 2002-07-15

```

```

; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: Patentin 3.1
; SEQ ID NO 8
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-10-020-510-8

```

Alignment Scores:	
Pred. No.:	5,77e-96
Score:	1118.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	96.80%
DB:	24
Length:	205
Matches:	205
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-899-303A-7 (1-633) x US-10-020-510-8 (1-209)

QY	1	ATGTGGGTAAAGGTCATCGATGATACCCCTTAAGCGGGCTTCGGCCAGCCATGAGGGTAAATT	60
Db	1	MettLeuGlyValIleAspThrLeuThnCysSylIyheAlaAspMetGlyTyrIle	20
QY	61	CCGCTCGTGGCGCCCCCTTAGGGGGGTCTCCAGAGCCCTGGCGCATGGCGTCCGGGTT	120
Db	21	ProLeuValGlyAlaProLeuGlyGlyAlaAlaIArgAlaLeuAlaHisGlyValArgVal	40
QY	121	CTGGAAGCGGGCGGCACTATGCAACAAGGAAATTGGCTGGTTCCTTTCTATCTTC	180
Db	41	LeuGlnAspGlyValAsnTyrAlaThrGlyAsnLeuProGlyCysSerPheSerIlePhe	60
QY	181	CTTTGGCTTACATGTCCTGCTGACCAATTCACAGCTTCGCGCTTATGAGTGGCCAAAGTG	240
Db	61	LeuLeuAlaLeuLeuSerCysLeuThrIleIerolAserAlaTyrGlyValaIArgAsnVal	80
QY	241	TTCGGGATGATACCATGTGCAGAACAGACTGCTCCAACTCAAGCATTTGTATGAGCAGCG	300
Db	81	SerGlyMetTyrHisValThrAsnAspCysSerAsnSerIleValTyrGlyAlaAla	100
QY	301	GACATGATCAATGCACACCCCGGGTGGTGGCCGCTTGGGGAGAACACTGTTCGCC	360
Db	101	AspMetIleMetHisThrProGlyCysValaIProCysValaIArgGlnAsnAsnSerArg	120
QY	361	TGCTGGGTAGGCGTCACCCCCCAGCCCTGCAAGCTAGGAAGCCAGAGCTCCCACTAGACA	420
Db	121	CysTrpValaIleLeuThrProThrLeuAlaIaIArgAsnIaIaSerValProThrThr	140
QY	421	ATPAGACGCCACGTCGATTTCTGCTGGGGCGGCTCTTTCTGTTCCGCTATGTAAGTCG	480
Db	141	IleArgArgHisValaIAspIleLeuValaGlyAlaIaIaIaPheCysSerAlaMetTyrVal	160
QY	481	GGGGATCTCTCCGGATGTCGTCTTCCTGCTGCCAGATGTTACACATCTGCGCCGGCGG	540
Db	161	GlyAspLeuCysGlySerValaIpheLeuValaIserGlnLeuPheThrIleSerProAlaArg	180
QY	541	CATGAGACGATGCAGAGCATGCAATTTGCTCAATCTATCTCCGGCCACATMACAGGTCAACGT	600
Db	181	HisGlnThrValaIaIAsnPCysAsnCysSerIleTyrProGlnHisIleThrGlnHisArg	200
QY	601	ATGGCTTCGGATGATGATGTAACATCG	627
Db	201	MetIaItrpAspMetMetIaIaItrp	209

RESULT 8
US-10-128-590-88
; Sequence 88, Application US/10128590

```

; GENERAL INFORMATION:
; APPLICANT: Immunogenetics N.V.
; TITLE OF INVENTION: Core-glycosylated HCV envelope proteins
; FILE REFERENCE: 135 PCT
; CURRENT APPLICATION NUMBER: US/10/128,590
; CURRENT FILING DATE: 2002-07-22
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: Patentin version 3.1

```


Db 141 IleArgArgHisValAspLeuLeuValGlyAlaAlaAlaPheCysSerAlaMetTyrVal 160
QY 481 GGGATCTCTGGGATCTGTCTTCCTGTCCTCCAGCTGTTCACCATCTCCGTCGGCG 540
Db 161 GlyAspLeuGlySerValPheLeuValSerGlnLeuPheThrIleSerProArgArg 180
QY 541 CATGAGACGGCTGCAGCACTGCATTCCTCAATTCATCCCGGCGCACATPAACAGTCCAGT 600
Db 181 HisGluThrValGlnAspCysAsnCysSerIleTyrProGlyHisIleThrGlyHisArg 200
QY 601 ATGGCTGGGATATGATGATGACTGG 627
Db 201 MetAlaTrpAspMetMetMetAsnTrp 209
RESULT 10
US-09-899-303A-6
Sequence 6, Application US/09899303A
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
BOSMAN, FONS
DE MARTYNOFF, GUY
BUYSE, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/899,303A
FILING DATE: 06-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/612,973
FILING DATE: 11-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 263 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-899-303A-6
Alignment Scores:
Pred. No.: 1,79e-95 Length: 263
Score: 1113.00 Matches: 207
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.04% Mismatches: 0
Query Match: 96.36% Indels: 0
Gaps: 0
US-09-899-303A-7 (1-633) x US-09-899-303A-6 (1-263)
QY 1 ATGTGGTAAAGTCATCATACCTTACGTCGGCTTCGGCGACATCATGAGGTACATT 60
Db 1 MetLeuGlyValIleAspTrpHisLeuThrLeuThrCysGlyPheAlaAspLeuValGlyTyrIle 20

QY 61 CCGCTCGCGCGCCCCCTAGGGGGTGTCCAGACCCCTGGGCGCATGGCGTTCGGGTT 120
Db 21 ProLeuValGlyAlaProLeuGlyGlyAlaAlaArgAlaLeuAlaHisGlyValArgVal 40
QY 121 CTGGAACAGCGCGTGAACATATGACACAGGGAATTTGGCTGGTCTTTCTATCTTC 180
Db 41 LeuGluAspGlyValAsnTyrAlaThrGlyAsnLeuProGlyCysSerPheSerIlePhe 60
QY 181 CTCTTGGCTTACTCTGCTCTGTTCACCATTCACACTTCGGTTATAGAGTCCGAAGTG 240
Db 61 LeuLeuAlaLeuLeuSerCysLeuThrValProAlaSerAlaTyrGluValAlaAsnVal 80
QY 241 TCCGGATGTACATGTCACAGAACACTGCTCCAACTCAACATTCATGTATGAGAGCAGCG 300
Db 81 SerGlyMetTyrHisValThrAsnAspCysSerAsnSerIleValTyrGluAlaAla 100
QY 301 GACATGATCATGCACACCCCGGGGTGCGCTGCTGCGGAGAACATCTTCCGCG 360
Db 101 AspMetIleMetHisThrProGlyCysValProCysValArgGluAsnAsnSerSerArg 120
QY 361 TGCCTGGTACGCGTCACCCCGGCTGCGAGCTAGAACCGCACCGTCCCACTACGACA 420
Db 121 CysTrpAlaAlaLeuThrProThrLeuAlaAlaArgAsnAlaSerValProThrThrThr 140
QY 421 ATACGACGCCACGTCGATTGCTGCTGGGGCGGCTCTTCTGCTCCGTTATGTACGTG 480
Db 141 IleArgArgHisValAspLeuLeuValGlyAlaAlaAlaPheCysSerAlaMetTyrVal 160
QY 481 GGGATCTCTGGGATCTGTCTTCCTGTCCTCCAGCTGTTCACCATCTCCGTCGGCG 540
Db 161 GlyAspLeuGlySerValPheLeuValSerGlnLeuPheThrIleSerProArgArg 180
QY 541 CATGAGACGGCTGCAGCACTGCATTCCTCAATTCATCCCGGCGCACATPAACAGTCCAGT 600
Db 181 HisGluThrValGlnAspCysAsnCysSerIleTyrProGlyHisIleThrGlyHisArg 200
QY 601 ATGGCTGGGATATGATGATGACTGG 627
Db 201 MetAlaTrpAspMetMetMetAsnTrp 209
RESULT 11
US-09-973-025-6
Sequence 6, Application US/09973025
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
BOSMAN, FONS
DE MARTYNOFF, GUY
BUYSE, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/973,025
FILING DATE: 10-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/612,973
FILING DATE: 11-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.

1	REGISTRATION NUMBER: 32,205
2	REFERENCE/DOCKET NUMBER: 1487-10
3	TELECOMMUNICATION INFORMATION:
4	TELEPHONE: (703) 816-4000
5	TELEFAX: (703) 816-4100
6	INFORMATION FOR SEQ ID NO: 6:
7	SEQUENCE CHARACTERISTICS:
8	LENGTH: 263 amino acids
9	TYPE: amino acid
10	TOPOLOGY: linear
11	MOLECULE TYPE: protein
12	SEQUENCE DESCRIPTION: SEQ ID NO: 6:
13	US-09-973-025-6
14	Alignment Scores:
15	Pred. NO.: 1,79e-95
16	Score: 1113.00
17	Percent Similarity: 100.00%
18	Best Local Similarity: 99.04%
19	Query Match: 96.36%
20	DB: 23
21	Gaps: 0
22	US-09-899-303A-7 (1-633) x US-09-973-025-6 (1-263)
23	QY 1 ATGTTGGGTAAAGTCATCGATGATACCTTACGTGGGGCTTGCCGACCTCATGGGGTACATT 60
24	Db 1 MetLeuGlyLysValIleAspThrLeuThrCysGlyPheAlaAspLeuValGlyTyrIle 20
25	QY 61 CCGGTCGCGGGCGGGCCCGCTTGGGGGGGTCGCCAGAGCCCGGGCGCATGGCGGCCGCTT 120
26	Db 21 ProLeuValGlyAlaProLeuGlyGlyAlaAlaIleArgAlaLeuAlaHisGlyValArgVal 40
27	QY 121 CTGGAAGACGGCGTGAACCTATGCAACAAGGAAATTTGGCTGGTCTTTCTGTATCTTC 180
28	Db 41 LeuGluAspGlyValAsnTyrAlaThrGlyAsnLeuProGlyCysSerPheSerIlePhe 60
29	QY 181 CTCCTGGGTACTACGTCCTGTCTGACCATTCACCTCCGCTTACCGCTTATGAGTGGCAAGTG 240
30	Db 61 LeuLeuAlaLeuLeuSerCysLeuThrValProAlaSerIleTyrGluValArgVal 80
31	QY 241 TCCGGGATGTACCATGTCTCAGCAAGACAGCTGCTCAACTCAAGCAATTTGTATGAGCGACG 300
32	Db 81 SerGlyMetTyrHisValThrAsnAspCysSerAsnSerIleValTyrGluAlaAla 100
33	QY 301 GACATGATCATGACACACCCCGGGTGGCTGCCCTTCGCTGGCGGAGAACAATCTTCCCGC 360
34	Db 101 AspMetIleMetHisThrProGlyCysValProCysValArgGluAsnAsnSerSerArg 120
35	QY 361 TGCCTGGGATGAGGTCACCCCCAGCCTCGACAGCTGTGAAGCGCAGCGCTCCACTAGACA 420
36	Db 121 CysTrpValAlaLeuThrProThrLeuAlaIleArgAsnAlaSerValProThrThrThr 140
37	QY 421 ATACGACGCCACGTGATTTGGCTGTTGGGGCGGCTGCTTTCTGTTCGCCGTATGTACGTG 480
38	Db 141 IleArgArgHisValAspLeuLeuValGlyAlaAlaIlePheCysSerAlaMetTyrVal 160
39	QY 481 GGGGATCTCTCCGGGATGCTGTCTTCCCTGCTGCCACAGCTGTTCACACATCTGGCTCGCCG 540
40	Db 161 GlyAspLeuCysGlySerValPheLeuValSerIleuPheThrIleSerProArgArg 180
41	QY 541 CATGAGACGTGCGAGCATGTCAATTTGCTCATCTATATCCCGGACCATATAAGGTACACGT 600
42	Db 181 HisGluThrValGlnAspCysAsnCysSerIleTyrProGlyHisIleThrGlyHisArg 200
43	QY 601 ATGGCTTGGATATGATGATGACTGG 627
44	Db 201 MetaIleArgAspMetMetMetAsnTrp 209
45	RESULT 12
46	US-09-995-791-6
47	Sequence 6, Application US/09995791
48	GENERAL INFORMATION:
49	APPLICANT: Innogenetics N.V.

```

; TITLE OF INVENTION: Purified hepatitis C virus envelope proteins for diagnostic use.
; TITLE OF INVENTION: Therapeutic use.
; FILE REFERENCE: 2551-68
; CURRENT APPLICATION NUMBER: US/09/995,791
; CURRENT FILING DATE: 2001-11-29
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn 3.1
; SEQ ID NO 6
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-995-791-6

Alignment Scores:
Pred. No.: 1,79e-95 Length: 263
Score: 1113.00 Matches: 207
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.04% Mismatch: 0
Query Match: 96.36% Indels: 0
DB: 23 Gaps: 0

US-09-899-303A-7 (1-633) x US-09-995-791-6 (1-263)

QY 1 ATGTGGGTAAAGTCATGATGATACCTTACGTGCGGCTTCCGCCACATCATAGGGTACATT 60
Db 1 MetLeuGlyLysValIleAspThrLeuThrCysGlyPheAlaAspLeuValGlyTyrIle 20
QY 61 CCGGTCGTCGCGGCCCCCTAGGGGTGCTGTCGCACAGAGCCATGGCCATAGGGGTCCGGGTT 120
Db 21 ProLeuValGlyAlaProLeuGlyGlyAlaAlaArgAlaLeuAlaHisGlyValArgVal 40
QY 121 CTGAAGACGCGCGGTACATATGCAACAGAGGAATTTGCCGTGTGCTCTTTCTATCTTC 180
Db 41 LeuGluAspGlyValAsnTyrAlaThrGlyLysLeuProGlyCysSerPheSerIlePhe 60
QY 181 CTCCTGGCTTTACGTGCTCGTCTGACATTCGACATTCGAGTTCGCGTTATGAGTGCAGACTG 240
Db 61 LeuLeuAlaLeuLeuSerCysLeuThrValProAlaSerAlaTyrGlyValArgAsnVal 80
QY 241 TCCGGGATGTACCATATGTCACGACGACGCTGCCAATCACTCAACATTTGTATGAGCAGCG 300
Db 81 SerGlyMetLysValIleThrAsnAspCysSerAsnSerIleValTyrGlyAlaIle 100
QY 301 GACATGATCATGACACACCCCGGGGTGCGTGGCTTCGCTTGGGAGACAACTCTTCCCGC 360
Db 101 AspMetIleMetHisIsthProGlyCysValProCysValArgGluAsnAsnSerArg 120
QY 361 TGCCTGGTAGGCGCTACACCCCGACGCTGCAGGTAGGAGGACGCGAGGCTCCCACTAGACA 420
Db 121 CysThrValAlaLeuThrProThrLeuAlaAlaArgAsnAlaSerValProThrThr 140
QY 421 ATACAGCCCAACGTCGATTTGCTCGTGGGCGCGCTTCTGTTCGCTATGACGTG 480
Db 141 IleArgArgHisValAspLeuLeuValGlyAlaAlaAlaPheCysSerAlaMetLysVal 160
QY 481 GGGGATTCCTGGGATTCGTCTCTCTGCTCCGCCAGCTGTACACATGTGCGCTCGCGG 540
Db 161 GlyAspLeuCysGlySerValPheLeuValSerGlnLeuPheThrIleSerProArgArg 180
QY 541 CATGAGACGGTGCAGAGCTGCATATGCTCAATCTTCCCGGACACATTAACAGTCAACCGT 600
Db 181 HisGluThrValGlnAspCysAsnGlySerIleTyrProGlyHisIleThrGlyHisArg 200
QY 601 ATGCGTTGGGATATGATGATGACTGG 627
Db 201 MetAlaTyrAspMetMetMetAsnTyr 209

RESULT 13
US-09-995-808-6
; Sequence 6, Application US/09995808
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: Purified hepatitis C virus envelope proteins for diagnostic use.

```

```
; TITLE OF INVENTION: therapeutic use.
; FILE REFERENCE: 2531-70
; CURRENT APPLICATION NUMBER: US/09/995,808
; CURRENT FILING DATE: 2001-11-29
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn 3.1
; SEQ ID NO 6
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-995-808-6

Alignment Scores:
Pred. No.: 1.79e-95 Length: 263
Score: 1113.00 Matches: 207
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.04% Mismatches: 0
Query Match: 96.36% Indels: 0
DB: 23 Gaps: 0

US-09-899-303a-7 (1-633) x US-09-995-808-6 (1-263)
QY 1 ATGTTGGGTAAAGTCATCGATACCTTACGTCGGCGCTTCGCCGACCTCATGGGGTACATT 60
Db 1 MetLeuGlyysValIleAspThrLeuThrCysGlyPheAlaAspLeuValGlyTyrIle 20
QY 61 CCGCTCTCGCGCCGCCCTAGAGGGGTGTGCCAGAGCCCTGGCGCATGGCGCTCCGGGTT 120
Db 21 ProLeuValGlyAlaProLeuGlyGlyAlaAlaArgAlaLeuAlaHisGlyValArgVal 40
QY 121 CTGGAGACGGCGGTGAACATATGCAACAGGAAATTTGCCCTGGTTCCTTTCTATCTTC 180
Db 41 LeuGluAspGlyValAsnTyrAlaThrGlyAsnLeuProGlyCysSerPheSerIlePhe 60
QY 181 CTCTTGCTTACTGCTGCTGTGACCATTCACCTCCGCTTACGCTTATGAGTGGCAAGTG 240
Db 61 LeuLeuAlaLeuLeuSerCysLeuThrValProAlaSerAlaTyrGluValArgAsnVal 80
QY 241 TCCGGATGTACCATGTACAGAACAGACTGCTCAACTCAAGCATTTGTATGAGGCAGCG 300
Db 81 SerGlyMetTyrHisValThrAsnAspCysSerAsnSerSerIleValTyrGluAlaIle 100
QY 301 GACATGATCATGCACACCCCGGGGTGGTCCCTGCGCTTGGGGAACAACCTTCCCGC 360
Db 101 AspMetIleMetHisThrProGlyCysValProCysValArgGluAsnAsnSerSerArg 120
QY 361 TGCCTGGGTAGCGGTACCCCGACGCTGCGAGCTGGAACGCGACGCTCCCACTAGCACA 420
Db 121 CysThrPValAlaLeuThrProThrLeuAlaIleArgAsnAlaSerValProThrThrThr 140
QY 421 ATACGACGCCACGTGATTTGCTGCTGGGGGCGGCTCTTCTGTCTCCGATATGACGTG 480
Db 141 IleArgThrHisValAspLeuLeuValGlyAlaAlaIlePheCysSerAlaMetTyrVal 160
QY 481 GGGGATCTTCGCGATGTGCTTCCTGCTGCTCCAGCTGTTCACCAATCGCCTCGCCGG 540
Db 161 G1yAspLeuCysGlySerValPheLeuValSerGlnLeuPheThrIleSerProArgArg 180
QY 541 CATGAGACGGGTGCGAGCTGCAATTGCTCAATCATCCGGGCCCATAAACGCGCACGCT 600
Db 181 HisGluThrValGlnAspCysAsnCysSerIleTyrProGlyHisIleThrGlyHisArg 200
QY 601 ATGGCTTGGATATGATGATGAACCTGG 627
Db 201 MetAlaTrpAspMetMetMetAlaSerTrp 209

RESULT 14
US-09-995-860-6
; Sequence 6, Application US/09995860
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: Purified hepatitis C virus envelope proteins for diagnostic and
; TITLE OF INVENTION: therapeutic use.
```

```
; FILE REFERENCE: 2531-69
; CURRENT APPLICATION NUMBER: US/09/995,860
; CURRENT FILING DATE: 2001-11-29
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn 3.1
; SEQ ID NO 6
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-995-860-6

Alignment Scores:
Pred. No.: 1.79e-95 Length: 263
Score: 1113.00 Matches: 207
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.04% Mismatches: 0
Query Match: 96.36% Indels: 0
DB: 23 Gaps: 0

US-09-899-303a-7 (1-633) x US-09-995-860-6 (1-263)
QY 1 ATGTTGGGTAAAGTCATCGATACCTTACGTCGGCGCTTCGCCGACCTCATGGGGTACATT 60
Db 1 MetLeuGlyysValIleAspThrLeuThrCysGlyPheAlaAspLeuValGlyTyrIle 20
QY 61 CCGCTCTCGCGCCGCCCTAGAGGGGTGTGCCAGAGCCCTGGCGCATGGCGCTCCGGGTT 120
Db 21 ProLeuValGlyAlaProLeuGlyGlyAlaAlaArgAlaLeuAlaHisGlyValArgVal 40
QY 121 CTGGAGACGGCGGTGAACATATGCAACAGGAAATTTGCCCTGGTTCCTTTCTATCTTC 180
Db 41 LeuGluAspGlyValAsnTyrAlaThrGlyAsnLeuProGlyCysSerPheSerIlePhe 60
QY 181 CTCTTGCTTACTGCTGCTGTGACCATTCACCTCCGCTTACGCTTATGAGTGGCAAGTG 240
Db 61 LeuLeuAlaLeuLeuSerCysLeuThrValProAlaSerAlaTyrGluValArgAsnVal 80
QY 241 TCCGGATGTACCATGTACAGAACAGACTGCTCAACTCAAGCATTTGTATGAGGCAGCG 300
Db 81 SerGlyMetTyrHisValThrAsnAspCysSerAsnSerSerIleValTyrGluAlaIle 100
QY 301 GACATGATCATGCACACCCCGGGGTGGTCCCTGCGCTTGGGGAACAACCTTCCCGC 360
Db 101 AspMetIleMetHisThrProGlyCysValProCysValArgGluAsnAsnSerSerArg 120
QY 361 TGCCTGGGTAGCGGTACCCCGACGCTGCGAGCTGGAACGCGACGCTCCCACTAGCACA 420
Db 121 CysThrPValAlaLeuThrProThrLeuAlaIleArgAsnAlaSerValProThrThrThr 140
QY 421 ATACGACGCCACGTGATTTGCTGCTGGGGGCGGCTCTTCTGTCTCCGATATGACGTG 480
Db 141 IleArgThrHisValAspLeuLeuValGlyAlaAlaIlePheCysSerAlaMetTyrVal 160
QY 481 GGGGATCTTCGCGATGTGCTTCCTGCTGCTCCAGCTGTTCACCAATCGCCTCGCCGG 540
Db 161 G1yAspLeuCysGlySerValPheLeuValSerGlnLeuPheThrIleSerProArgArg 180
QY 541 CATGAGACGGGTGCGAGCTGCAATTGCTCAATCATCCGGGCCCATAAACGCGCACGCT 600
Db 181 HisGluThrValGlnAspCysAsnCysSerIleTyrProGlyHisIleThrGlyHisArg 200
QY 601 ATGGCTTGGATATGATGATGAACCTGG 627
Db 201 MetAlaTrpAspMetMetMetAlaSerTrp 209

RESULT 15
US-10-020-510-6
; Sequence 6, Application US/10020510
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: Purified hepatitis C virus envelope proteins for diagnostic an
; TITLE OF INVENTION: therapeutic use.
; FILE REFERENCE: 2531-72
```

; CURRENT APPLICATION NUMBER: US/10/020,510
; CURRENT FILING DATE: 2002-07-15
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn 3.1
; SEQ ID NO 6
; LENGTH: 263
; TYPE: PRP
; ORGANISM: Hepatitis C virus
US-10-020-510-6

Alignment Scores:

Pred. No.:	1,79e-95	length:	263
Score:	1113.00	Matches:	207
Percent Similarity:	100.00%	Conservative:	2
Best Local Similarity:	99.04%	Mismatches:	0
Query Match:	96.36%	Indels:	0
DB:	24	Gaps:	0

US-09-899-303a-7 (1-633) x US-10-020-510-6 (1-263)

QY 1 ATGTTGGGTAAAGGTCATTCGATACCTTAAGTGGGCTGCGCCGACCTCATGGGGTACATT 60
|||||
Db 1 MetLeuGlyLyValIleAspThrLeuThrCysGlyPheAlaAspLeuValGlyTyrIle 20
61 CCGCTGCTGGCGCCCGCCCTAGGGGGTCTGCCAGAGCCCTGGCCATGGCGTCCGGGTT 120
|||||
Db 21 ProLeuValGlyAlaProLeuGlyGlyAlaAlaArgAlaLeuAlaHisGlyValArgVal 40
121 CTGGAAGACGGCGGTGAATGCAACAGGGAATTTGCTGCTGCTCTTCTCTATCTTC 180
|||||
Db 41 LeuGluAspGlyValAsnTyrAlaThrGlyAsnLeuProGlyCysSerPheSerIlePhe 60
181 CTCTTGGCTTTACTGCTGCTGCTGACCAATTCAGCTTCGCTTATGAGTGGCAACGTG 240
|||||
Db 61 LeuLeuAlaLeuLeuSerCysLeuThrValProAlaSerAlaTyrGluValArgAsnVal 80
241 TCCGGGATGTACCATGTCACAGACGACTGCTCCAACTCAAGCATTTGTATGAGGACGG 300
|||||
Db 81 SerGlyMetTyrHisValIleThrAsnAspCysSerAsnSerIleValIleTyrGluAlaAla 100
301 GACATGATCATGCACACCCCGGGGCTGCGCTGCGGTCGGGAGAACAACCTTCCCGC 360
|||||
Db 101 AspMetIleMetHisThrProGlyCysValProCysValArgGluAsnAsnSerSerArg 120
361 TGCTGGGTAGCGCTCAACCCCGACGCTCGACGAGTAGGAACGCCAGCGTCCCACTACGACA 420
|||||
Db 121 CysTrpValAlaLeuThrProThrLeuAlaAlaArgAsnAlaSerValProThrThrThr 140
421 ATACGACGCCACGTCGATTTGCTGCTGGGGCGGCTGCTTCTGTTCCGCTATGTACGTG 480
|||||
Db 141 IleArgArgHisValAspLeuLeuValGlyAlaAlaAlaPheCysSerAlaMetTyrVal 160
481 GGGGATCTGCGGATCTGCTCTCTGCTGCTCCACGCTGTCACGATCTCGCCTCGCCGG 540
|||||
Db 161 GlyAspLeuCysGlySerValPheLeuValSerGlnLeuPheThrIleSerProAlaArg 180
541 CATGACAGCGGTGACGAGCTGCAATTCCTCAATCATCCGGCCACATAACAGGTACCGGT 600
|||||
Db 181 HisGluThrValGlnAspCysAsnCysSerIleTyrProGlyHisIleThrGlyHisArg 200
601 ATGGCTTGGGATATGATGATGAAGTGG 627
|||||
Db 201 MetAlaIlePaspMetMetMetAsnTrp 209

Search completed: February 12, 2003, 09:08:57
Job time : 245 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 12, 2003, 08:55:00 ; Search time 61 Seconds
(without alignments)
4276.322 Million cell updates/sec

Title: US-09-899-303a-7
Perfect score: 1155
Sequence: 1 ATGTTGGGTAGGTCATCGA.....TGATGATGAGACGTGTAATAG 633

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 1343160

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL-frame+ n2p.model -DEV-xip
-O-cgnt2.1/uspro.apool/us09899303/runat.12022003.085449.28268/app.query.fasta.1.775
-DB-SPRMBL.21 -OFMT-fastan -SUFFIX-n2p.rspt -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS-bits -START=1 -END=1 -MATRIX-biosum62 -TRANS-human40.cdl
-LIST=45 -DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE-LOCAL -OUTFMT=ptc -NORM-text -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER-US09899303 -CGCN.1.1.88 -eunat.12022003.085449.28268 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-MAIN.TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database :

SPRMBL.21:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.podent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.bacteriaph:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result Query Match Length DB ID Description
1 1095 94.8 310 12 Q90048 Q90048 hepatitis c

2	1094	94.7	513	12	081221	081221 hepatitis c
3	1094	94.7	3010	12	002829	002829 h genome po
4	1092	94.5	3010	12	09WMX2	09WMX2 hepatitis c
5	1091	94.5	3010	12	09J3H2	09J3H2 hepatitis c
6	1090	94.4	3010	12	09OP06	09OP06 hepatitis c
7	1089	94.3	3010	12	09J310	09J310 hepatitis c
8	1089	94.3	3010	12	09J3G7	09J3G7 hepatitis c
9	1089	94.0	3010	12	09OJY3	09OJY3 hepatitis c
10	1086	94.0	3010	12	081760	081760 hepatitis c
11	1085	93.9	3010	12	092972	092972 hepatitis c
12	1085	93.9	3010	12	09OJY4	09OJY4 hepatitis c
13	1084	93.9	441	12	081776	081776 hepatitis c
14	1084	93.9	802	12	068520	068520 hepatitis c
15	1084	93.9	3010	12	09J3H7	09J3H7 hepatitis c
16	1084	93.9	3010	12	09J3H7	09J3H7 hepatitis c
17	1083	93.8	587	12	068518	068518 hepatitis c
18	1083	93.8	787	12	008244	008244 h genome po
19	1083	93.8	1008	12	089156	089156 hepatitis c
20	1083	93.8	3010	12	09D7D6	09D7D6 hepatitis c
21	1083	93.8	3013	12	09OJY0	09OJY0 hepatitis c
22	1083	93.8	3010	12	09OJY9	09OJY9 hepatitis c
23	1082	93.7	3010	12	09D7D9	09D7D9 hepatitis c
24	1082	93.7	3010	12	09J3H9	09J3H9 hepatitis c
25	1082	93.7	3010	12	09J3H5	09J3H5 hepatitis c
26	1082	93.7	3010	12	08OJL8	08OJL8 hepatitis c
27	1081	93.6	337	12	086880	086880 hepatitis c
28	1081	93.6	2864	12	09WJL8	09WJL8 hepatitis c
29	1081	93.6	2864	12	092974	092974 hepatitis c
30	1081	93.6	3010	12	08V638	08V638 hepatitis c
31	1081	93.6	3010	12	09D7E4	09D7E4 hepatitis c
32	1081	93.6	3010	12	092969	092969 hepatitis c
33	1081	93.6	3010	12	09OJY8	09OJY8 hepatitis c
34	1081	93.6	3010	12	09OJY7	09OJY7 hepatitis c
35	1081	93.6	3010	12	09OJX8	09OJX8 hepatitis c
36	1081	93.6	3010	12	09OJX7	09OJX7 hepatitis c
37	1081	93.6	3010	12	09J1A0	09J1A0 hepatitis c
38	1081	93.6	3010	12	09J1V3	09J1V3 hepatitis c
39	1080	93.5	2864	12	09WJL9	09WJL9 hepatitis c
40	1080	93.5	3010	12	092971	092971 hepatitis c
41	1080	93.5	3010	12	002828	002828 h genome po
42	1080	93.5	3012	12	09WJL7	09WJL7 hepatitis c
43	1079	93.4	1008	12	089157	089157 hepatitis c
44	1079	93.4	2864	12	092975	092975 hepatitis c
45	1079	93.4	3008	12	09J3F4	09J3F4 hepatitis c

ALIGNMENTS

RESULT 1
Q90048 PRELIMINARY: PRT: 310 AA.
ID Q90048
AC Q90048;
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-NOV-1996 (TREMblrel. 20, Last annotation update)
DE 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Genome polypeptide [Contains: envelope glycoprotein E2 (GP68) (GP70)
(NS1)] (Fragment).
GN POLYPEPTIDE.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_Taxid=11103;
RN [1]
RP MEDLINE=93019054; PubMed=1328487;
RA Kohara M., Tsukiyama-Kohara K., Maki N., Asano K., Yamaguchi K.,
Miki K., Tanaka S., Hattori N., Matsura Y., Saito I., et al;
RT Expression and characterization of glycoprotein gp35 of hepatitis C
virus using recombinant vaccinia virus.;
RL J. Gen. Virol. 73:2313-2318(1992).
CC -I- SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.
DR EMBL: S46012; AAB23552.1; -;
InterPro: IPR002521; HCV_core.

162 (67)

DR InterPro: IPR002519; HCV_env.
 DR InterPro: IPR002531; HCV_NSI.
 DR Pfam: PF01542; HCV_core; 1.
 DR Pfam: PF01539; HCV_env; 1.
 DR Pfam: PF01560; HCV_NSI; 1.
 DR ProDom: PD186062; HCV_NSI; 1.
 DR Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
 KM Polyprotein; Transmembrane.
 FT NON_TER 1
 SQ SEQUENCE 310 AA; 33200 MW; 5DB04CE7F70E41F9 CRC64;

Alignment Scores:
 Pred. No.: 2,35e-97 Length: 310
 Score: 1095.00 Matches: 203
 Percent Similarity: 98.568 Conservative: 2
 Best Local Similarity: 97.604 Mismatches: 3
 Query Match: 94.81% Indels: 0
 DB: 12 Gaps: 0

US-09-899-303A-7 (1-633) x Q90048 (1-310)

QY 4 TTGGGTAAAGTCATGATACCTTACGTCGGCTTCGCCACCTCATGGGGTACATTCCG 63
 |||||
 Db 3 TGGGlyValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIlePro 22
 |||||

QY 64 CTGCTGGGGCCCCCTAGGGGGTGTGCGCAGAGCCCTGGCGATGGCGGTTCTG 123
 |||||
 Db 23 TGGGlyValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIlePro 42
 |||||

QY 124 GAAGAGGGGGTGAATGATGACAGAGGAAATTTGCTGCTGCTTCTTATCTCCCTC 183
 |||||
 Db 43 GUAAGGlyValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIlePro 62
 |||||

QY 184 TTGGCTTAACTGTCGTGTGACCATTTCCAGCTTCGCTTATGAGTGGCGCAAGTCTCC 243
 |||||
 Db 63 TGGGlyValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIlePro 82
 |||||

QY 244 GGGATGTACATGTCAGAGAGAGCTGCTCAACTCAAGCATTTGTATGAGCGAGGAC 303
 |||||
 Db 83 GYIleTyrHisValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIlePro 102
 |||||

QY 304 ATGATGATGACAG 363
 |||||
 Db 103 MetIleMetHisThrProGlyCysValProGlyValIleArgGlyAsnSerSerAlaGlyCys 122
 |||||

QY 364 TGGGTAGCGGTCAACCCCGAGCTGCGAGCTGAGGAGAGAGAGAGAGAGAGAGAGAG 423
 |||||
 Db 123 TTPAlaIleLeuThrProThrLeuAlaAlaArgAsnThrSerValProThrThrIle 142
 |||||

QY 424 CGAGCGCAGCTGATTTGCTGCTGGGGGGGCTTTCGTTCCGCTATGATGAGTGGGG 483
 |||||
 Db 143 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 162
 |||||

QY 484 GATCTGGGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 543
 |||||
 Db 163 AspLeuCysGlySerValPheLeuValSerGlnLeuPheThrPheSerProAlaGlyHis 182
 |||||

QY 544 GAGAGCGTGCAG 603
 |||||
 Db 183 GATThValIleGlnAspCysSerIleTyrProGlyHisLeuThrGlyHisAlaGlyMet 202
 |||||

QY 604 GCTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 627
 |||||
 Db 203 AlATrPaspMetMetMetAsnTrp 210
 |||||

RESULT 2
 081221
 ID 081221 PRELIMINARY: PRT: 513 AA.
 AC 081221;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-BEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Genome polyprotein [contains: envelope glycoprotein E2 (GP68) (GP70)]

DE (NS1) (Fragment).
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OC NCBI_TaxID-11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-HC-04;
 RX MEDLINE-91013116; PubMed-2170712;
 RA Okamoto H., Okada S., Sugiyama Y., Yotsunoto S., Tanaka T.,
 RA Yoshizawa H.;
 RT "The 5'-terminal sequence of the hepatitis C virus genome.";
 RL Jpn. J. Exp. Med. 60:167-177(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-HC-04;
 RX MEDLINE-92044440; PubMed-1658196;
 RA Okamoto H., Okada S., Sugiyama Y., Kurai K., Iizuka H., Machida A.,
 RA Miyakawa T., Mayumi M.;
 RT "Nucleotide sequences of the genomic RNA of hepatitis C virus isolated
 RT from a human carrier: comparison with reported isolates for conserved
 RT and divergent regions.";
 RL J. Gen. Virol. 72:2697-2704(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-HC-04;
 RX MEDLINE-94270990; PubMed-7545932;
 RA Hotta H., Doi H., Hayashi T., Purwanta M., Soemarto W., Mizokami M.,
 RA Ohba K., Homma M.;
 RT "Analysis of the core and E1 envelope region sequences of a novel
 RT variant of hepatitis C virus obtained in Indonesia.";
 RL Arch. Virol. 136:53-62(1994).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-HC-04;
 RX MEDLINE-94174722; PubMed-7510436;
 RA Mink M., Benichou S., Kadaule P., Tiollais P., Prince A.,
 RA Inchausti G.;
 RT "Characterization and mapping of a B-cell immunogenic domain in
 RT hepatitis C virus E2 glycoprotein using a yeast peptide library.";
 RL Virology 200:246-255(1994).
 CC -1 SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.
 DR EMBL: D00832; BAA00706.1;
 DR InterPro: IPR002532; HCV_capsid.
 DR InterPro: IPR002521; HCV_core.
 DR InterPro: IPR002519; HCV_env.
 DR InterPro: IPR002531; HCV_NSI.
 DR Pfam: PF01543; HCV_capsid; 1.
 DR Pfam: PF01542; HCV_core; 1.
 DR Pfam: PF01539; HCV_env; 1.
 DR Pfam: PF01560; HCV_NSI; 1.
 DR ProDom: PD186062; HCV_NSI; 1.
 DR Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
 KM Polyprotein; Transmembrane.
 FT NON_TER 513
 SQ SEQUENCE 513 AA; 55874 MW; 3C8E1D94FED1FA7E2 CRC64;

Alignment Scores:
 Pred. No.: 3.15e-97 Length: 513
 Score: 1094.00 Matches: 203
 Percent Similarity: 99.528 Conservative: 4
 Best Local Similarity: 97.604 Mismatches: 1
 Query Match: 94.72% Indels: 0
 DB: 12 Gaps: 0

US-09-899-303A-7 (1-633) x Q81221 (1-513)

QY 4 TTGGGTAAAGTCATGATACCTTACGTCGGCTTCGCCACCTCATGGGGTACATTCCG 63
 |||||
 Db 119 TGGGlyValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIlePro 138
 |||||

QY 64 CTGCTGGGGCCCCCTAGGGGGTGTGCGCAGAGAGCCCTGGCGATGGCGGTTCTG 123
 |||||

```

Db 139 LeuValAlGlyAlaProLeuGlyGlyAlaAlaAlaGlyAlaLeuAlaHisGlyValAlaArgValLeu 158
Qy 124 GAAGAGCGCGGTGAATATGCAACAGGAAATTTGCTGTTCTTCTATCTTCCTC 183
Db 159 GluAspGlyValAsnThrGlyAlaSerGlyAsnLeuProGlyCysSerPheSerIlePheLeu 178
Qy 184 TTGGCTTAACTGCTGCTGTGACCAATTCAGCTTCCGCTTATGAGGTGGCAAGCTGCTC 243
Db 179 LeuAlaLeuLeuSerCysLeuThrIleProAlaSerAlaTyrGlyValAlaArgAsnValSer 198
Qy 244 GGGATGATACATGTCAGAACAGCACTGCTGCAACTGCAACTGATGTATGAGAGCGGAGC 303
Db 199 GlyIleTyrHisValThrAsnAspCysSerAsnSerSerIleValTyrGlyAlaAlaAsp 218
Qy 304 ATGATCATGCAACAGCGCGGTGCGGCTGCTGCGGAGAAACAATCTTCCGCTGCTC 363
Db 219 MetIleMetHisThrProGlyCysValProCysValAlaArgGluAspAsnSerSerArgCys 238
Qy 364 TTGGGTAGCGCTACAGCCCAAGCTGCGAGCTAGAGAACAGCGCTGCCACATGACAGATA 423
Db 239 TrpValAlaLeuThrProThrLeuAlaAlaArgAsnAlaSerValProThrThrIle 258
Qy 424 GCAGCGCACTGATGATGCTGCTGCGGGGCGCTGCTTCTGCTCCCTATGATGAGTGGGG 483
Db 259 ArgArgHisValAlaAspLeuValGlyAlaAlaAlaAlaPheCysSerAlaMetTyrValGly 278
Qy 484 GATCTCTGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 543
Db 279 AspLeuCysGlySerValAlaPheLeuValSerGlnLeuPheThrPheSerProAlaArgHis 298
Qy 544 GAGAGCGTGCAGAGTGCATATGCTCATATCCGCGCCACATAGAGTCAACGCTATG 603
Db 299 GluThrValAlaGlnAspCysAsnGlySerIleTyrProGlyHisLeuSerGlyHisArgMet 318
Qy 604 GCTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 627
Db 319 AlaTrpAspMetMetMetMetMet 326

RESULT 3
ID 002829 PRELIMINARY: PRT: 3010 AA.
AC 002829:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Genome polypeptide [CONTAINS: capsid protein C (Core protein): matrix protein NS1; nonstructural protein NS2; major envelope protein E; nonstructural protein NS1; nonstructural protein NS4B; helicase (NS3); RNA-directed RNA polymerase (EC 2.7.7.48) (RNA nucleotidyltransferase (RNA-directed)).]
DE Hepatitis C virus.
OS Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_Taxid=11103;
RN NCBI_Sequence=11103;
RP SEQUENCE FROM N.A.
RC STRAIN=HC-J4;
RX MEDLINE=92391112; PubMed=1325713;
RA Muchmore E.E., Peterson D.A., Ito Y., Mishito S., Iizuka H., Tanaka T., Okamoto H., Kojima M., Okada S.I., Yoshizawa H.,
"Genetic drift of hepatitis C virus during an 8.2-year infection in a chimpanzee: variability and stability."
RT Virology 190:894-899(1992).
RL -I- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE -> N PYROPHOSPHATE +
CC RN(N).
DR EMBL: D13558; BA02756.1; -
DR HSRP: F26663; LXP.
DR MEROPS: S29.001; -
DR InterPro: IPR000345; CysC_heme_bind.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR002522; HCV_capsid.
DR InterPro: IPR002521; HCV_core.

```

```

DR InterPro: IPR002519; HCV_env.
DR InterPro: IPR002531; HCV_NS1.
DR InterPro: IPR002518; HCV_NS2.
DR InterPro: IPR004109; HCV_NS3.
DR InterPro: IPR000745; HCV_NS4.
DR InterPro: IPR001490; HCV_NS4b.
DR InterPro: IPR002868; HCV_NS5a.
DR InterPro: IPR002166; HCV_RdRp.
DR pfam: PF01543; HCV_capsid. 1.
DR pfam: PF01542; HCV_core. 1.
DR pfam: PF01539; HCV_env. 1.
DR pfam: PF01560; HCV_NS1. 1.
DR pfam: PF01538; HCV_NS2. 1.
DR pfam: PF02907; HCV_NS3. 1.
DR pfam: PF01006; HCV_NS4a. 1.
DR pfam: PF01001; HCV_NS4b. 1.
DR pfam: PF01506; HCV_NS5a. 1.
DR pfam: PF00998; HCV_RdRp. 1.
DR PRODOM: PD186062; HCV_NS1. 1.
DR PROSITE: PS00190; CYTOCHROME C; UNKNOWN.1.
DR TRANSFERASE: Polypeptide; Glycoprotein; RNA-directed RNA polymerase;
KW Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
KW Transmembrane; Nonstructural protein.
FT CHAIN 1 190
FT CHAIN 191 382
FT CHAIN 383 842
FT CHAIN 843 1119
FT CHAIN 1120 1728
FT CHAIN 1729 2126
FT CHAIN 2127 3010
FT TRANSMEM 347 369
FT DOMAIN 796 803
FT NP_BIND 1230 1237
FT CARBOHYD 196 196
FT CARBOHYD 209 209
FT CARBOHYD 234 234
FT CARBOHYD 305 305
FT CARBOHYD 417 417
FT CARBOHYD 423 423
FT CARBOHYD 430 430
FT CARBOHYD 448 448
FT CARBOHYD 532 532
FT CARBOHYD 556 556
FT CARBOHYD 645 645
FT CARBOHYD 2041 2041
FT CARBOHYD 2077 2077
FT CARBOHYD 2240 2240
FT CARBOHYD 2788 2788
SQ SEQUENCE 3010 AA; 326959 MW; 93D465526F3EADP8 CRC64;

Alignment Scores:
Pred. No.: 4,02e-97 Length: 3010
Score: 1094.00 Matches: 203
Percent Similarity: 99.52% Conservative: 4
Best Local Similarity: 97.60% Mismatches: 1
Query Match: 94.72% Indels: 0
DB: 12 Gaps: 0

US-09-899-303a-7 (1-633) x 002829 (1-3010)
Qy 4 TTGGGTAGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 63
Db 119 LeuGlyLysValAlaIleAspThrLeuThrCysGlyPheAlaAspLeuGlyTyrIlePro 138
Qy 64 CTCGTGCGGCGCCCTAGGGGCTGCTGCGAGAGCCCTGCGCATGCGCTGCGGTTCTG 123
Db 139 LeuValAlGlyAlaProLeuGlyGlyAlaAlaAlaAlaGlyAlaLeuAlaHisGlyValAlaArgValLeu 158
Qy 124 GAAGAGCGCGGTGAATATGCAACAGGAAATTTGCTGTTCTTCTATCTTCCTC 183
Db 159 GluAspGlyValAsnThrGlyAlaSerGlyAsnLeuProGlyCysSerPheSerIlePheLeu 178

```


QY	544	GAGAGGGGCGAGCATCAATGGTCATCATCTCCGGGACACATACAGGTCACCGATG	603
Db	239	GIUHVVAIGINASPQVSASVCYSSEITLIERPRGGLYHISVALINMGLYHISATGMC	318
QY	604	GCTTGCGATATGATGATGACACTCG	627
Db	319	Alatrypaspmetmetelantirp	326
RESULT 5			
ID	09J3H2	PREDIMINARY;	PRT; 3010 AA.
AC	09J3H2;		09J3H2;
DT	01-OCT-2000 (TREMBLrel. 15, Created)		
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)		
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
DE	Genome polypeptide [Contains: envelope glycoprotein E2 (GP68) (GP70) (NS1)].		
OS	Hepatitis C virus.		
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;		
OC	Hepadnavirus.		
OX	NCBI_Taxid=11103;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-M020;		
RA	Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.		
RT	*Characteristics of hepatitis C viral genome associated with disease		
RT	progression. *;		
RL	Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.		
CC	-1- SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.		
DR	EMBL; AF207761; AAF65951.1; -.		
DR	HSSP; P26663; INS3.		
DR	MEROPS; S29_001; -.		
DR	MEROPS; U39_001; -.		
DR	InterPro; IPR000345; CytC_heme_bind.		
DR	InterPro; IPR001410; DEAD.		
DR	InterPro; IPR002522; HCV_capsid.		
DR	InterPro; IPR002521; HCV_core.		
DR	InterPro; IPR002519; HCV_env.		
DR	InterPro; IPR002531; HCV_NS1.		
DR	InterPro; IPR002518; HCV_NS2.		
DR	InterPro; IPR004109; HCV_NS3.		
DR	InterPro; IPR000745; HCV_NS4.		
DR	InterPro; IPR001490; HCV_NS4b.		
DR	InterPro; IPR002868; HCV_NS5a.		
DR	InterPro; IPR002166; HCV_RDRP.		
DR	pfam; PF01543; HCV_capsid; 1.		
DR	pfam; PF01542; HCV_core; 1.		
DR	pfam; PF01539; HCV_env; 1.		
DR	pfam; PF01560; HCV_NS1; 1.		
DR	pfam; PF01538; HCV_NS2; 1.		
DR	pfam; PF02907; HCV_NS3; 1.		
DR	pfam; PF01006; HCV_NS4a; 1.		
DR	pfam; PF01001; HCV_NS4b; 1.		
DR	pfam; PF01506; HCV_NS5a; 1.		
DR	pfam; PF00998; HCV_RDRP; 1.		
DR	ProDom; PD186062; HCV_NS1; 1.		
DR	PROSITE; PS00190; CYTOCHROME C; UNKNOWN.1.		
KW	Cost protein, Envelope protein, Glycoprotein, Nonstructural protein,		
KW	polyprotein, RNA-directed RNA polymerase, Transmembrane		
QJ	SEQUENCE 3010 AA; 326763 MW; 1A48EE4B51440D0 CRC64;		

Alignment Scores:			
Pred. No.:	7,836-97	Length:	3010
Score:	1091.00	Matches:	201
Percent Similarity:	99.04%	Conservative:	5
Best Local Similarity:	96.63%	Mismatches:	2
Query Match:	94.66%	Indels:	0
Dbs:	12	Gaps:	0
US-09-899-303A-7 (1-633) x G9J3H2 (1-3010)			
4 TTGGGTAAGGCTATCGATACCCCTTACGTGCGGCTCCGCGACCTCATGAGGGTACATTCG			633

Dd		119	Lcnciluyxvalliileasphricleuthircysglrphenalaarblemetgltgtlliepro	13
Oy		64	CTGTCGGCCCCCCTAGGGGTGCSTGCCAGAGCCSTGGCGCATGGCGTGGGTTTCG	12
Dd		139	Lcuvalcllyalabroeuicglylalaalaargalaleualanlsilyvaladgvaalleu	15
Oy		124	GAAACAGCGGTGAACATATGCACAACAGGGAATTGGCTGGTGCTCTTCTATCTTCCTC	18
Dd		159	GLUuspolylalamnrylatnrlatnrglyslansleuproglycysserphenetlliepheleu	17
Oy		184	TTCGCTTTACTGTCTGTCTGCASCATTCSCAGCTCCSGCTTAATGAGGTGCCAAGCTCCS	24
Dd		179	Lcuallaleuleusercysleuthrlieproliaseratalatylgltvaladgaanvalser	19
Oy		244	GGGATGTACCATGTCAACAGAACACATGCTTCSAACSTADACATTTGTGTAGAGCGAGAC	30
Dd		199	GLYValIygrlnavalthrasmnpsycssersenalaserllevaltylgltlaalaalabar	21
Oy		304	ATGATCATGCACACCCCCGGGTGGCGCCCTCGTTCCGGAGAACATCTTCCCGCTGC	36
Dd		219	Mettlemetisthrproglcysvalprocyvaladgluaanaanserleratdycs	23
Oy		364	TGGGTAGCGCTCACCCCCAGCTCGCAGCTAGGAACGGCAGCGTCCCCTACGACATA	42
Dd		239	Trrvalialaleuthrprothrleuilaalaadgaanalaserlleprothrlnrhlrval	25
Oy		424	CGAGCGCCAGCATTTGGCGTGGGGGGCGTGGTCTTCTCGCTTAGTACATGCGGGG	48
Dd		259	Atryamrtastalsaspheuleuvalglualaalaalarnecysseralametgtvalglu	27
Oy		484	GATCTGTGGGATCTGTCTTCTCTGCTCCAGCTGTTCACACATCTGCGCTCGCGGAT	54
Dd		279	Aspleucysglyseralrphenleuvalserglnleupherlrpheserprodargnlis	29
Oy		544	GAGAGCGGTGAGAGATGCAANTGCTCAATCTATCCCGGGSACATAAAGTCAACGCTATG	60
Dd		299	GIuthrtvalidmnpysasnrcysserilletyrproglynlasvaltlnrghlnlasargmet	31
Oy		604	GCTTGGGATGATGATGATGAACGTG	62
Dd		319	Alatrpsmethemetslntrp	32
RESULT 6				
OQ9P06				
ID	OQ9P06	PRELIMINARY;	PRT; 3010 AA.	
AC	OQ9P06:			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	Genome polyprotein [contains: envelope glycoprotein E2 (GP68) (GP70) (NS1)].			
OS	Hepatitis C virus type 1b.			
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;			
OX	Hepadnavirus.			
NCBI_TaxID=31647;				
RN	[1]			
RC	SEQUENCE FROM N.A.			
RP	STRAIN-RB;			
RA	Bartenschlager R.;			
RL	Submitted (MAY-1999) to the EMBL/Genbank/DBJ databases.			
RN	[2]			
RC	SEQUENCE FROM N.A.			
RP	STRAIN-RB;			
RA	MEDLINE=99370154; PubMed=10438800;			
RT	Koch J.O., Bartenschlager R.;			
RT	"Modulation of hepatitis C virus NS5A hyperphosphorylation by			
RT	nonstructural proteins NS3, NS4A, and NS4B."			
RL	J. Virol. 73:7138-7146(1999).			
CC	-1. SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.			
DR	EMBL; AJ238800; CAB53095.1; -.			
DR	HSP; P2663; INS3.			
DR	MEROPS; S29_001; -.			

DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR002522; HCV_capsid.
 DR InterPro: IPR002521; HCV_core.
 DR InterPro: IPR002519; HCV_env.
 DR InterPro: IPR002531; HCV_NS1.
 DR InterPro: IPR002518; HCV_NS2.
 DR InterPro: IPR004109; HCV_NS3.
 DR InterPro: IPR000745; HCV_NS4.
 DR InterPro: IPR001490; HCV_NS4b.
 DR InterPro: IPR002868; HCV_NS5a.
 DR InterPro: IPR002166; HCV_RdRp.
 DR pfam: PF01543; HCV_capsid; 1.
 DR pfam: PF01542; HCV_core; 1.
 DR pfam: PF01539; HCV_env; 1.
 DR pfam: PF01560; HCV_NS1; 1.
 DR pfam: PF01538; HCV_NS2; 1.
 DR pfam: PF02907; HCV_NS3; 1.
 DR pfam: PF01006; HCV_NS4a; 1.
 DR pfam: PF01001; HCV_NS4b; 1.
 DR pfam: PF01506; HCV_NS5a; 1.
 DR pfam: PF00998; HCV_RdRp; 1.
 DR ProDom: PD186062; HCV_NS1; 1.
 DR ATP-binding: Coat protein; Envelope protein; Glycoprotein; Helicase;
 KM Nonstructural protein; Polyprotein; RNA-directed RNA polymerase;
 KM Transmembrane.
 FT CHAIN 1 191 CORE PROTEIN.
 FT CHAIN 192 383 GLYCOPROTEIN E1.
 FT CHAIN 384 746 GLYCOPROTEIN E2.
 FT CHAIN 747 809 P1 PEPTIDE.
 FT CHAIN 810 1026 NS2 PROTEINASE.
 FT CHAIN 1027 1657 NS3 PROTEINASE/HELICASE.
 FT CHAIN 1658 1711 NS3/4A PROTEINASE COFACTOR.
 FT CHAIN 1712 1972 NS4B PROTEIN.
 FT CHAIN 1973 2418 NS5A PHOSPHOPROTEIN.
 FT CHAIN 2419 3010 NS5B RNA DEPENDENT RNA POLYMERASE.
 SQ SEQUENCE 3010 AA; 326999 MW; A570BB980DD64634 CRC64;

Alignment Scores:
 Pred. No.: 9.79e-97 Length: 3010
 Score: 1090.00 Matches: 201
 Percent Similarity: 99.04% Conservative: 5
 Best Local Similarity: 96.63% Mismatches: 2
 Query Match: 94.37% Indels: 0
 DB: 12 Gaps: 0

US-09-899-303A-7 (1-633) x Q90P06 (1-3010)

QY 4 TTGGGTAAAGTATCGATACCCCTTACGTGGCGGCTTCGACCTCATGGGTCATTCGCG 63
 DB 119 LKUGLYLYVALILLESPIHRLHNCYSGLYPHEALASPHEUWETGLYTYTLEUPRO 138
 QY 64 CTCGTGGGCCCCCTAGGGGGTGTGCCAGAGCCCTGGCGATGGCGTCCGGTTCTG 123
 DB 139 LKVALIGLYALPHEUENGLYALALALARGALALEUALHISGLYVALARGVALLEU 158
 QY 124 GAAGAGCGGCTGAACATATGACACAGGGAATTCCTGCTTCCTATCTTCCTC 183
 DB 159 GLUASPGLYVALASNTYRLATHRGLYASNLEUPROGLYCYSERHESERTLEPHELEU 178
 QY 184 TTGGCTTACTGCTGCTGTGACCATTCAGCTTCGCGTTATGAGGTGGCAACGTCGC 243
 DB 179 LKALALEUENLSEICYSLEUTHRIELEPROLASERATLYGLVALALRGSNVALSER 198
 QY 244 GGGATGATCATGTCAAGNAGCACTGCTCAACTCAAGCATTTGTATGAGGCGGAGC 303
 DB 199 GLYVALTYRHISVALTRHNSPSCYSSEASERANALASERILEVALYRGLUALALASAP 218
 QY 304 ATGATATGACACACCCCGGGTGGCGCTTCGCTGGGAGAACACATCTTCCCGCTGC 363
 DB 219 METILEMERHISHPROGLYCYSVALIPROCYVALARGGLUASNSERISERATGYS 238
 QY 364 TGGGTAGCGCTACCCCGCGCGCTGAGTAGAAGCGCAGGCTCCCACTACGACAATA 423

DB 239 TRPVALLALEUHPROTHRLLEUALALARGASNALISERVALPROTHRTHRILE 258
 QY 424 CGAGCCACGTCGATTTGCTGTTGGGGCGGCTTCTGTTCCGCTATGATGAGGGG 483
 DB 259 ARGARGHISVALASPLEUENVALIGLYALALALALEUCYSSERIALMETTYRVALISGLY 278
 QY 484 GATCTCTGCGGATCTGTTCTCTGCTGCTCCACAGCTGTTCACCATTCGCGCTCGCGGCAT 543
 DB 279 ASPLEUCYSGLYSERVALPHELEUVALALAGLNEUPHETHPHESEPROARGHIS 298
 QY 544 GAGAGCGTGACGACGTCGATTCCTCAATCTATCCGGCCACATACAGTACCGCTATG 603
 DB 299 GLUTHRALGLHNASPCYSANCYSSERIETLYRPROGLYHISVALTRGLYHISARGMET 318
 QY 604 GCTTGGATATGATGATGACTGG 627
 DB 319 ALATRPASMETMETMETASNTRP 326
 RESULT 7
 ID Q9J310 PRELIMINARY; PRT: 3010 AA.
 AC Q9J310;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Genome polyprotein [Contains: envelope glycoprotein E2 (GP68) (GP70) (NS1)].
 DE Hepatitis C virus.
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 NCBI_Taxid=11103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MD12;
 RA Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.;
 RT "Characteristics of hepatitis C viral genome associated with disease progression."
 RT Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.
 DR EMBL: AF207753; AAF65943.1; -.
 DR HSSP: P26663; LUXP.
 DR MEROPS: S29.001; -.
 DR MEROPS: U39.001; -.
 DR InterPro: IPR000345; Cyto_heme_bind.
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR002522; HCV_capsid.
 DR InterPro: IPR002521; HCV_core.
 DR InterPro: IPR002519; HCV_env.
 DR InterPro: IPR002531; HCV_NS1.
 DR InterPro: IPR002518; HCV_NS2.
 DR InterPro: IPR004109; HCV_NS3.
 DR InterPro: IPR000745; HCV_NS4.
 DR InterPro: IPR001490; HCV_NS4b.
 DR InterPro: IPR002868; HCV_NS5a.
 DR InterPro: IPR002166; HCV_RdRp.
 DR InterPro: IPR001650; Helicase_C.
 DR pfam: PF01543; HCV_capsid; 1.
 DR pfam: PF01542; HCV_core; 1.
 DR pfam: PF01539; HCV_env; 1.
 DR pfam: PF01560; HCV_NS1; 1.
 DR pfam: PF01538; HCV_NS2; 1.
 DR pfam: PF02907; HCV_NS3; 1.
 DR pfam: PF01006; HCV_NS4a; 1.
 DR pfam: PF01001; HCV_NS4b; 1.
 DR pfam: PF01506; HCV_NS5a; 1.
 DR pfam: PF00998; HCV_RdRp; 1.
 DR pfam: PF00271; Helicase_C; 1.
 DR ProDom: PD186062; HCV_NS1; 1.
 DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN.1.
 DR Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
 KM Polyprotein; RNA-directed RNA polymerase; Transmembrane.
 SQ SEQUENCE 3010 AA; 326692 MW; 074098DB305AFA1A9 CRC64;

Alignment Scores:

Pred. No.: 1,22e-96 Length: 3010
 Score: 1089.00 Matches: 201
 Percent Similarity: 98.08% Conservative: 3
 Best Local Similarity: 96.63% Mismatches: 4
 Query Match: 94.29% Indels: 0
 DB: 12 Gaps: 0

US-09-899-303a-7 (1-633) x Q9J3G7 (1-3010)

```

QY 4 TTGGGTAAAGTCAATGATACCTTACGTGGGCGCTTGGCCGACCTCATGGGGTACATTCCG 63
    |||||||
Db 119 LeuGlyLysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIlePro 138
QY 64 CTCGTGGCGCCGCCCTAGAGGGGTGCTCCAGAGCCCTGGCGCATGGCGGTCTGTG 123
    |||||||
Db 139 LeuValGlyAlaProLeuGlyGlyAlaAlaArgValLeuAlaHisGlyValArgValLeu 158
QY 124 GAAGACGGCGGTGAACATATGCAACAGGGAATTTGCTGTGTTCTCTATCTTCCTC 183
    |||||||
Db 159 GluAspGlyValAsnTyrAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeu 178
QY 184 TTGGCTTACTGTCCTGTGACCATTCAGCTTCGCTATGAGTGCGGCAACGTGTCC 243
    |||||||
Db 179 LeuAlaLeuLeuSerCysLeuThrThrProAlaSerAlaTyrGlyValArgAsnValSer 198
QY 244 GGGATGATACCATGTCAGAGAGCACTGCTCCAGCTCAAGCATTTGTATGAGCGAGCGAC 303
    |||||||
Db 199 GlyMetCysTrpIleValThrAsnAspCysSerAsnSerSerIleValTyrGlyAlaAlaAsp 218
QY 304 ATGATCATGCAACACCCCGGGGTGCGTCCCTGCTGGGAGAAACAATCTTCCCGCTGC 363
    |||||||
Db 219 MetIleMetHisAlaProGlyCysValProCysValArgGlyAsnAsnSerSerArgCys 238
QY 364 TTGGTAGCGGTCAACCCCGACGTGCGAGCTAGGAAGCGAGCGGCCCTAGAGCAATA 423
    |||||||
Db 239 TrpValAlaLeuThrProThrLeuAlaAlaArgAsnAlaSerValProThrThrIle 258
QY 424 CGAGCGCACGTGATTTGCTGCTTGGGCGCGCTGCTTCTGTTCCGCTATGTAAGTGGG 483
    |||||||
Db 259 ArgArgHisValAspLeuLeuValGlyAlaAlaAlaPheCysSerAlaMetTyrValGly 278
QY 484 GATCTGCGCGGATCTGCTTCTGCTGCTGCTCCAGCTGTCAACATTCGCGCTCCGCGGAT 543
    |||||||
Db 279 AspLeuCysGlySerValPheLeuIleSerGlyLeuPheThrPheSerProArgArgHis 298
QY 544 GAGCGGTGAGGACATGCTATGCTCAATGATCCCGCGGACCAATACAGGTCCACCTATG 603
    |||||||
Db 299 GluThrValGlnAspCysAsnCysSerIleTyrProGlyHisValSerGlyHisAlaGlyMet 318
QY 604 GCTTGGGATATGATGATGAATGCG 627
    |||||||
Db 319 AlaTrpAspMetMetMetAsnTrp 326

RESULT 8
Q9J3G7 PRELIMINARY: PRT: 3010 AA.
AC Q9J3G7:
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Genome polyprotein [Contains: envelope glycoprotein E2 (GP68) (GP70)
DE (NS1)].
OS Hepatitis C virus.
OC Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=MD25;
RA Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.;
RT "Characteristics of hepatitis C viral genome associated with disease
    progression.";
  
```

RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 CC -I- SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.
 DR EMBL: AF207766; AAF65956.1; -
 DR HSP: P26663; LUXP.
 DR MEROPS: S29.001; -
 DR MEROPS: U39.001; -
 DR InterPro: IPR000345; Cytic_heme_bind.
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR002522; HCV_capsid.
 DR InterPro: IPR002521; HCV_core.
 DR InterPro: IPR002519; HCV_env.
 DR InterPro: IPR002531; HCV_NS1.
 DR InterPro: IPR002518; HCV_NS2.
 DR InterPro: IPR004109; HCV_NS3.
 DR InterPro: IPR000745; HCV_NS4a.
 DR InterPro: IPR001490; HCV_NS4b.
 DR InterPro: IPR002868; HCV_NS5a.
 DR InterPro: IPR002166; HCV_RdRp.
 DR InterPro: IPR001650; Helicase_C.
 DR Pfam: PF01543; HCV_capsid; 1.
 DR Pfam: PF01539; HCV_env; 1.
 DR Pfam: PF01560; HCV_NS1; 1.
 DR Pfam: PF01538; HCV_NS2; 1.
 DR Pfam: PF02907; HCV_NS3; 1.
 DR Pfam: PF01006; HCV_NS4a; 1.
 DR Pfam: PF01001; HCV_NS4b; 1.
 DR Pfam: PF01506; HCV_NS5a; 1.
 DR Pfam: PF00998; HCV_RdRp; 1.
 DR Pfam: PF00271; Helicase_C; 1.
 DR ProDom: PD186062; HCV_NS1; 1.
 DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
 KW Coat protein: Envelope protein; Glycoprotein; Nonstructural protein;
 KW polyprotein; RNA-directed RNA polymerase; Transmembrane.
 SQ SEQUENCE 3010 AA; 326507 MW; 9E8AC8ADA4B44DF CRC64;

Alignment Scores:

Pred. No.: 1,22e-96 Length: 3010
 Score: 1089.00 Matches: 202
 Percent Similarity: 98.56% Conservative: 3
 Best Local Similarity: 97.12% Mismatches: 3
 Query Match: 94.29% Indels: 0
 DB: 12 Gaps: 0

US-09-899-303a-7 (1-633) x Q9J3G7 (1-3010)

```

QY 4 TTGGGTAAAGTCAATGATACCTTACGTGGGCGCTTGGCCGACCTCATGGGGTACATTCCG 63
    |||||||
Db 119 LeuGlyLysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIlePro 138
QY 64 CTCGTGGCGCCGCCCTAGAGGGGTGCTCCAGAGCCCTGGCGCATGGCGGTCTGTG 123
    |||||||
Db 139 LeuValGlyAlaProLeuGlyGlyAlaAlaArgValLeuAlaHisGlyValArgValLeu 158
QY 124 GAAGACGGCGGTGAACATATGCAACAGGGAATTTGCTGTGTTCTCTATCTTCCTC 183
    |||||||
Db 159 GluAspGlyValAsnTyrAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeu 198
QY 179 LeuAlaLeuLeuSerCysLeuThrThrProAlaSerAlaTyrGlyValArgAsnValSer 218
QY 244 GGGATGATACCATGTCAGAGAGCACTGCTCCAGCTCAAGCATTTGTATGAGCGAGCGAC 303
    |||||||
Db 199 GlyValTyrHisValThrAsnAspCysSerAsnSerSerIleValTyrGlyAlaAlaAsp 218
QY 304 ATGATCATGCAACACCCCGGGGTGCGTCCCTGCTGGGAGAAACAATCTTCCCGCTGC 363
    |||||||
Db 219 MetIleMetHisAlaProGlyCysValProCysValArgGlyAsnAsnSerSerArgCys 238
QY 364 TTGGTAGCGGTCAACCCCGACGTGCGAGCTAGGAAGCGAGCGGCCCTAGAGCAATA 423
    |||||||
Db 239 TrpValAlaLeuThrProThrLeuAlaAlaArgAsnAlaSerValProThrThrIle 258
  
```

```

OY 424 CGAGCCACAGTCGATTGCTGCTGGGGGCGGCTGCTTCTGTTCGGATATGATGGGG 483
    |||||||
Db 239 ArgAGHISValAspLeuValGlyAlaAlaAlaPheCysSerAlaMetGlyValGly 278
OY 484 GATCTCGGGGATCTGCTCTCTCCGCCAGCTGTCACACATCGCTCGGCGGGCAT 543
    |||||||
Db 279 AspLeuGlySerValPheLeuValSerGlnLeuPheThrPheSerProArgArgHis 298
OY 544 GAGACGGTCAGAGCAATGCTCAATCTATCCCGGCCACATGAACAGGTCACCGTATG 603
    |||||||
Db 299 GluThrValGlnAspCysAsnGlySerIleTyrProGlnHisLeuSerGlnHisArgMet 318
OY 604 GCTGGGATATGATGATGAATGCG 627
    |||||||
Db 319 AlaTrpAspMetMetMetAsnTrp 326

RESULT 9
OQ01Y3 PRELIMINARY: PRT: 3010 AA.
AC 0901Y3;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Genome polyprotein [Contains: envelope glycoprotein E2 (GP68) (GP70) (NS1)].
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Md-2;
RX MEDLINE=20013325; PubMed=10544098;
RA Nagayama K., Kurosaki M., Enomoto N., Maekawa S.Y., Miyasaka Y.,
  Tarawa J.I., Izumi N., Marumo F., Sato C.;
RT "Time-related changes in full-length hepatitis C virus and hepatitis
  C activity.";
RL Virology 263:244-253(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Md-2;
RA Nagayama K., Kurosaki M., Enomoto N., Maekawa S., Miyasaka Y.,
  Sakamoto N., Fukuma T., Tarawa J., Izumi N., Marumo F., Sato C.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1 SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.
DR EMBL: AF165052; AAD56187.1; -.
DR HSSP: P26663; 1UXP.
DR MEROPS: S29.001; -.
DR MEROPS: U39.001; -.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR002522; HCV_capsid.
DR InterPro: IPR002521; HCV_core.
DR InterPro: IPR002519; HCV_env.
DR InterPro: IPR002531; HCV_NS1.
DR InterPro: IPR002518; HCV_NS2.
DR InterPro: IPR004109; HCV_NS3.
DR InterPro: IPR000745; HCV_NS4.
DR InterPro: IPR001490; HCV_NS4b.
DR InterPro: IPR002868; HCV_NS5a.
DR InterPro: IPR002166; HCV_RdRp.
DR Pfam: PF01543; HCV_capsid; 1.
DR Pfam: PF01542; HCV_core; 1.
DR Pfam: PF01539; HCV_env; 1.
DR Pfam: PF01560; HCV_NS1; 1.
DR Pfam: PF01538; HCV_NS2; 1.
DR Pfam: PF02907; HCV_NS3; 1.
DR Pfam: PF01006; HCV_NS4a; 1.
DR Pfam: PF01001; HCV_NS4b; 1.
DR Pfam: PF01506; HCV_NS5a; 1.
DR Pfam: PF00998; HCV_RdRp; 1.
DR ProDom: PD186062; HCV_NS1; 1.
KW ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase;

```

```

KM Nonstructural protein; Polyprotein; RNA-directed RNA polymerase;
KW Transmembrane.
SQ SEQUENCE 3010 AA; 327087 MW; 2EF78321686D4002 CRC64;

Alignment Scores:
Pred. No.: 2,39e-96 Length: 3010
Score: 1086.00 Matches: 200
Percent Similarity: 99.04% Conservative: 6
Best Local Similarity: 96.15% Mismatches: 2
Query Match: 94.03% Indels: 0
DB: 12 Gaps: 0

US-09-899-303a-7 (1-633) x OQ01Y3 (1-3010)
OY 4 TTGGTAAAGTCATGATGATACCTTACGTGCGGCTTCCGCGACCTGATGGGATATCCG 63
    |||||||
Db 119 LeuGlyLysValIleAspThrIleuThrGlySerPheAlaAspLeuMetGlyTyrIlePro 138
OY 64 CTGTCGGGCGCCCGCCCTAGGGGCTGTCGACAGCCCTGGGSCATGGGCTCGGCTGTCG 123
    |||||||
Db 139 LeuValGlyAlaProLeuGlyGlyAlaAlaArgAlaLeuAlaHisGlyValArgValLeu 158
OY 124 GAACAGCGCGTAACTATAGCAACAGGAAATTTGCGTGGCTGTTCTGTATGCTCTGC 183
    |||||||
Db 139 GluAspGlyValAsnTyrAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeu 178
OY 184 TTGGCTTACGTCTGCTGTCGACATTCGACGCTTCCGCTTATGAGTGGGACGTCGCC 243
    |||||||
Db 179 LeuAlaLeuLeuSerCysLeuThrIleProAlaSerAlaTyrGlyValArgAsnValSer 198
OY 244 GGGATGACCATGTCAGACAGACAGCTGTCGCAATCAGACATTTGTTATGAGCAGCGGAC 303
    |||||||
Db 199 GlyIleTyrHisValThrAsnAspCysSerAsnSerIleValTyrGlyAlaAlaAsp 218
OY 304 ATGACGACGACACACCCCGGGGTCGTCGCGCTGTCGGGAGAAACACTTCCCGCTGC 363
    |||||||
Db 219 ValIleMetHisThrProGlyCysValProCysValArgGlySerAsnSerIleGlyCys 238
OY 364 TTGGTAGCGCTCACCCCGCCAGCGCTCGACGTAGGAAGCCAGCGCTCCCACTACGACATA 423
    |||||||
Db 239 TrpValAlaLeuThrProThrIleuAlaAlaArgAlaAlaSerValProThrThrThrIle 258
OY 424 CGAGCCACAGTCGATTGCTGCTGGGGGCGGCTGCTTCTGTTCGGATATGATGGGG 483
    |||||||
Db 239 ArgAGHISValAspLeuValGlyThrAlaAlaAlaPheCysSerAlaMetGlyValGly 278
OY 484 GATCTCGGGGATCTGCTCTCTCCGCCAGCTGTCACACATCGCTCGGCGGGCAT 543
    |||||||
Db 279 AspLeuGlySerValPheLeuValSerGlnLeuPheThrPheSerProArgArgHis 298
OY 544 GAGACGGTCAGAGCAATGCTCAATCTATCCCGGCCACATGAACAGGTCACCGTATG 603
    |||||||
Db 299 GluThrValGlnAspCysAsnGlySerIleTyrProGlnHisLeuSerGlnHisArgMet 318
OY 604 GCTGGGATATGATGATGAATGCG 627
    |||||||
Db 319 AlaTrpAspMetMetMetAsnTrp 326

RESULT 10
O81760 PRELIMINARY: PRT: 3010 AA.
AC 081760;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Genome polyprotein [Contains: envelope glycoprotein E2 (GP68) (GP70) (NS1)].
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.

```


KM	Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW	Polyprotein; Transmembrane.
FT	NON_TER 441 441
SQ	SEQUENCE 441 AA; 47875 MW; 78025F07957E41E CRC64;
 Alignment Scores:	
Pred. No.:	2.86e-96 Length: 441
Score:	1084.00 Matches: 199
Percent Similarity:	98.56% Conservative: 3
Best Local Similarity:	95.67% Mismatches: 6
Query Match:	93.85% Indels: 0
DB:	Gaps: 0
 US-09-899-303A-7 (1-633) x 081776 (1-441)	
QY	4 TTGGGTAAAGTCATCGATACCATCCCTTACGCGGCGCTTCGCCGACGCATAGGGGTACATTCCG 63
Db	119 LeugllylsvalilIeaSPrlrLeuthrCyselylphelaInslapleumetGIyrrilepro 138
QY	64 CTCGTCGGCCCCCCTTAGGGGGTGCTGCCAGACGCCCTGGCGCATGGCGCGGTTCTG 123
Db	139 LeuvnlglIalArploLenglylAlalAlargrAlaleuAlahslgIyValArgrValleu 158
QY	124 GAAGAAGCGCGTGAACATATGACAAGGGGAATTGGCTGGTGGCTGTCTTCTATCTTC 183
Db	159 GlusarpGlYvalAsnTYrAlAlThrClYAsnleuProclYsserPheSerllepheleu 178
QY	184 TTGGCTTTACTGTCCTGTCTGACCATTCCAGCTTCGCGCTTATGAGGTGGCAAGTGTCC 243
Db	179 LeuAlaleuLeuSerCysIeuThrIlleProAlaserAlatyGrIuValArgrsnValser 198
QY	244 GGGATGTACCAATPGCACAGACAGTGGCTCCAACATCAAGCAATPTGTATAGGACGGGAC 303
Db	199 GlylleryrhIsvalThlnAsnAPCyserIsnsersetlerlevallyGrIualAlalaSp 218
QY	304 ATGATCATGCAACACCCCGGGTGGCGCTCCCTTCGGGAGAAACATCTTCCCGCTGC 363
Db	219 ValIlmetHslAlaproGLlCYuvalProcyvalArgrIunshnsersetlerArgys 238
QY	364 TGGGTAAGCGCTCACCCCACGCTGGCAAGTtNGMAAGCCAGGCTCCCACTACGACATA 423
Db	239 TrYvalAlaleuThrProthrleuAlalAlargrsnAlaserValProthrTrlrtrhu 258
QY	424 CGAGCCACAGTCGATTTGGCTGGTGGGGCGGCTCTTCTGTTCCGGTANGTACGTGGG 483
Db	259 ArgArghslsvalArpleuleuValGlYrhlAlalAlarheyserralameTyrrAlglY 278
QY	484 GATCTCTCGGAGATGTCTTCTCTGCTCCAGCTGTtTACCAATTCGCGCTCGCGGCA 543
Db	279 AsplreucysGIyserValrPheluIllerSerclInleuPherhrPheserProArgrAhIs 298
QY	544 GAGACGGTGCAGAGCTGAATGTGTCAATATTCGCCGACATPAACAGGTACGCTAG 603
Db	299 GIunrhrValglhsnAPCyssAsnCysSerllerTyProclYhIsvalSerlglyhIsArget 318
QY	604 GCTTGGATATGATGAGAACGCG 627
Db	319 AlArtpasmetMetelastITTP 326
 PRELIMINARY; PRT; 802 AA.	
ID	068520
AC	068520;
DT	01-NOV-1996 (TREMBLrel_01, Created)
DT	01-NOV-1996 (TREMBLrel_01, Last sequence update)
DT	01-DEC-2001 (TREMBLrel_19, Last annotation update)
DE	Genome polyprotein [Contains: envelope glycoprotein E2 (GP68) (GP70) (NS1)] (Fragment).
OS	Hepatitis C virus.
VN	Virusess; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC	Hepadnavirus.
OX	NCBI_TaxId=11103;
RN	[1]

```

RP  SEQUENCE FROM N.A.
RC  STRAIN-HTX.
RA  Mueller H.W., Selzer S., Heller A.E., Pfaff E., Theilmann L.;
RT  "Recurrent hepatitis C virus infection after liver transplantation in
RT  a patient negative for viral proteins: A possible role of immune
RT  tolerance.";
RU  Submitted (JAN-1996) to the EMBL/Genbank/DBJ databases.
CC  -1- SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.
DR  EMBL: U45463; AAA86919.1; -.
DR  MEROPS: U39.001; -.
DR  InterPro: IPR002522; HCV_capsid.
DR  InterPro: IPR002521; HCV_core.
DR  InterPro: IPR002519; HCV_env.
DR  InterPro: IPR002518; HCV_NS1.
DR  InterPro: IPR002518; HCV_NS2.
DR  Pfam: PF01543; HCV_capsid; 1.
DR  Pfam: PF01542; HCV_core; 1.
DR  Pfam: PF01539; HCV_env; 1.
DR  Pfam: PF01560; HCV_NS1; 1.
DR  Pfam: PF01538; HCV_NS2; 1.
DR  ProDom: PD186062; HCV_NS1; 1.
KW  Coat protein; Envelope protein; glycoprotein; Nonstructural protein;
KW  Polyprotein; Transmembrane.
FT  NON_TER
SQ  SEQUENCE 802 AA; 87900 MW; F17246CAAAA476 CRC64;

Alignment Scores:
Pred. No.: 3,11e-96 Length: 802
Score: 1084.00 Matches: 201
Percent Similarity: 99.04% Conservative: 5
Best Local Similarity: 96.63% Mismatches: 2
Query Match: 93.85% Indels: 0
DB: Gaps: 0

US-09-899-303A-7 (1-633) x Q68520 (1-802)
OY 4 TTGGGTAAGGTCATGATACCTTACGTCGGCTTCGCCGACCTCATGSGGTACATTCGG 63
DB 119 LeuGlyLysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetIlyTrIlePro 138
OY 64 CTGCTGCGGGCCCCCTAGGGGGTCTGCCAGAGCCCTGGCGCATGGCTCGGGGTTCTG 123
DB 139 LeuValGlyAlaProLeuGlyGlyAlaAlaAlaGlyAlaLeuAlaHisGlyAlaArgValIleu 158
OY 124 GAAGACGGCGTAACTATGCAACAGGGAATTTGCTGTTCTCTATCTTCTCTC 183
DB 159 GluAspGlyValAlaSerIlyThrGlyAsnLeuProGlyCysSerPheSerIlePheLeu 178
OY 184 TTGGGTTTACTGTCGTCGTCGATCCATTCACGTCCTCCGCTTATGAGTGGCGCAACGTCTCC 243
DB 179 LeuAlaLeuLeuSerCysLeuThrIleProAlaSerAlaIlyrGlyValAlaArgAsnValSer 198
OY 244 GGGATGACCATGATGACAGACAGTGTCTCAACTCAAGCATGTGTATGAGCAGCGGAC 303
DB 199 GlyValIlyrHisValIlyrHisAspCysSerAsnSerIleValIlyrIlyAlaAlaAsp 218
OY 304 ATGATCATGACACACCCCGGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 363
DB 219 MetIleMetHisThrProGlyCysValAlaProCysValAlaGlyIlyAsnAsnSerSerArgCys 238
OY 364 TGGGTAGCGCTCACCCGCCAGCTCCGAGTAGAAGACCCAGCGCTCCCGCATGACGACATA 423
DB 239 TrpValAlaLeuThrProThrLeuAlaAlaArgSerAlaAsnValSerThrThrIle 258
OY 424 CGACGCGACGTCGATTTGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 483
DB 259 ArgAlaGHisValAlaSerLeuValGlyAlaAlaAlaAlaPheCysSerAlaMetIlyrValGly 278
OY 484 GATCTCTGGGATCTGTCTTCTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 543
DB 279 AspLeuCysGlySerValAlaPheLeuValSerGlnLeuPheThrPheSerProAlaGHis 298
OY 544 GAGACGGTCGACGACTGCAATTGTCATCTATCCCGGCCACATAACAGGTACCCGTATG 603

```

```

DB 299 GluThrValGlnAspCysAsnCysSerIleIlyrProGlyHisValSerCylHisArgMet 318
OY 604 GCTTGGATATGATGATGACTGG 627
DB 319 AlaTrpAspMetMetCysIleTrp 326

RESULT 15
OY3H7 PRELIMINARY; PRT; 3010 AA.
AC OY3H7;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Genome polyprotein (Contains: envelope glycoprotein E2 (GP68) (GP70)
DE (NS1)).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MD15;
RA Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.;
RT "Characteristics of hepatitis C viral genome associated with disease
RT progression.";
RT Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.
DR EMBL: AF207756; AAF65946.1; -.
DR HSRP: P26663; IJXP.
DR MEROPS: S29.001; -.
DR MEROPS: U39.001; -.
DR InterPro: IPR000345; Cytochrome_b10d.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR002522; HCV_capsid.
DR InterPro: IPR002521; HCV_core.
DR InterPro: IPR002519; HCV_env.
DR InterPro: IPR002531; HCV_NS1.
DR InterPro: IPR002518; HCV_NS2.
DR InterPro: IPR004109; HCV_NS3.
DR InterPro: IPR000745; HCV_NS4a.
DR InterPro: IPR001490; HCV_NS4b.
DR InterPro: IPR002868; HCV_NS5a.
DR InterPro: IPR002166; HCV_RdRp.
DR Pfam: PF01543; HCV_capsid; 1.
DR Pfam: PF01542; HCV_core; 1.
DR Pfam: PF01539; HCV_env; 1.
DR Pfam: PF01560; HCV_NS1; 1.
DR Pfam: PF01538; HCV_NS2; 1.
DR Pfam: PF02907; HCV_NS3; 1.
DR Pfam: PF01006; HCV_NS4a; 1.
DR Pfam: PF01001; HCV_NS4b; 1.
DR Pfam: PF01506; HCV_NS5a; 1.
DR Pfam: PF00998; HCV_RdRp; 1.
DR ProDom: PD186062; HCV_NS1; 1.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
KW Coat protein; Envelope protein; glycoprotein; Nonstructural protein;
KW Polyprotein; RNA-directed RNA polymerase; Transmembrane.
SQ SEQUENCE 3010 AA; 327365 MW; D8653F7317FFA106 CRC64;

Alignment Scores:
Pred. No.: 3.72e-96 Length: 3010
Score: 1084.00 Matches: 200
Percent Similarity: 97.60% Conservative: 3
Best Local Similarity: 96.15% Mismatches: 5
Query Match: 93.85% Indels: 0
DB: Gaps: 0

US-09-899-303A-7 (1-633) x Q9J3H7 (1-3010)
OY 4 TTGGTAAAGGTCATGATACCTTACGTCGGCTTCGCCGACCTCATGSGGTACATTCGG 63
DB 119 LeuGlyLysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetIlyTrIlePro 138

```

```
QY 64 CTGCTGGGCCCCCTAGGGGGCTGTCAGAGCCCTGGCGCATGGCGTCCGGTTCTG 123
Db 139 LeuValGlyAlaProLeuGlyGlyAlaAlaArgAlaLeuAlaHisGlyValArgValLeu 158
QY 124 GAAGAGGGGCTGAACATGTAACAGGAAATTTGGCTGTTGCTCTCTATCTTCCTC 183
Db 159 GluAspGlyValAsnGlyAlaThrGlyAsnLeuProGlyCysSerPheSerLeuPheLeu 178
QY 184 TTGGCTTTACTGTCCTGTCGACCATTTCCAGCTTCCGCTATGAGTGGCCAACTGTC 243
Db 179 LeuAlaLeuLeuSerCysLeuThrProValSerAlaTyrGluValArgAsnValSer 198
QY 244 GGGATGTACCATGTACAGAAAGACTGCTCAACTCAAGCATTTGTATGAGGACGGAC 303
Db 199 GlyValTyrHisValThrAsnAspCysSerAsnSerSerLeuValTyrGluAlaAlaasp 218
QY 304 ATGATCATGTGACACCCCGGGGGTGGCTGCTGGCTGGGAGAACACTCTCCCGCTGC 363
Db 219 MetIleMetHisAlaProGlyCysValProCysValArgGluAsnAsnSerSerArgCys 238
QY 364 TGGGTAGCGCTCACCCCGACGCTCGCAGTAGGAACGCCAGCGTCCCGCTACGACAATA 423
Db 239 TrpValAlaLeuThrProThrLeuAlaAlaArgAsnIleSerIleProThrThrIle 258
QY 424 GAGAGCCACGTCGATTTGCTCGTTGGGGCGGCTGCTTCTGTCCGCTATGAGTGGGG 483
Db 259 ArgArgHisValAspLeuLeuValGlyAlaAlaAlaPheCysSerAlaMetTyrValGly 278
QY 484 GATCTCGCGGATGCTGCTCTCCGCTGCTCCAGCTGTTCAACATCTCGCCTCGCGGAT 543
Db 279 AspleucysGlySerValPheLeuValSerGlnLeuPheThrPheSerProArgArgHis 298
QY 544 GAGACGGTCGAGAGCTGCAATGCTCAATCTATCCCGGCACATACAGGTACACCGTATG 603
Db 299 GluThrValGlnAspCysAsnCysSerIleTyrProGlyHisValThrGlyHisArgMet 318
QY 604 GCTTGGGATATGATGATGAACCTGG 627
Db 319 AlaTrpAspMetMetMetAsnTrp 326
```

Search completed: February 12, 2003, 09:01:01
Job time : 79 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 18, 2003, 04:57:02 ; Search time 2140 Seconds

(without alignments)
8608.447 Million cell updates/sec

Title: US-09-899-303a-7

Perfect score: 633
Sequence: 1 ATGTGGGTAAAGTCATCGA.....TGATGATGAACGTGTAATAG 633

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_scs:*
28: em_un:*
29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pln:*
35: em_hcg_rod:*
36: em_hcg_man:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hcgo_hum:*
40: em_hcgo_mus:*
41: em_hcgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	633	100.0	633	6 A48669	A48669 Sequence 7
2	633	100.0	633	6 AR157326	AR157326 Sequence 7
3	633	100.0	633	6 AX452756	AX452756 Sequence
4	612.8	96.8	636	6 A48675	A48675 Sequence 13
5	612.8	96.8	636	6 AR157329	AR157329 Sequence
6	612.8	96.8	636	6 AX452762	AX452762 Sequence
7	608.8	96.2	795	6 A48667	A48667 Sequence 5
8	608.8	96.2	795	6 AR157325	AR157325 Sequence
9	608.8	96.2	795	6 AX452754	AX452754 Sequence
10	605.8	95.7	2082	6 A48709	A48709 Sequence 47
11	605.8	95.7	2082	6 AR157350	AR157350 Sequence
12	605.8	95.7	2082	6 AX452796	AX452796 Sequence
13	605.8	95.7	2433	6 A48711	A48711 Sequence 49
14	605.8	95.7	2433	6 AR157351	AR157351 Sequence
15	605.8	95.7	2433	6 AX452798	AX452798 Sequence
16	574.2	90.7	9379	14 AR207766	AR207766 Hepatitis
17	569.4	90.0	9386	14 AR165056	AR165056 Hepatitis
18	567.8	89.7	9379	14 AR165052	AR165052 Hepatitis
19	567.8	89.7	9379	14 HCVPOLYP	AJ000009 Hepatitis
20	567.8	89.7	9386	14 AR165055	AR165055 Hepatitis
21	566.2	89.4	9410	14 HPCCK1R2	D50481 Hepatitis C
22	564.6	89.2	1880	14 HPC5TRJ4	D00832 Hepatitis C
23	564.6	89.2	2540	6 E04260	E04260 CDNA encod1
24	564.6	89.2	2540	6 E04805	E04805 CDNA to 5'-
25	564.6	89.2	2540	6 E07391	E07391 CDNA encod1
26	564.6	89.2	9379	14 AR165047	AR165047 Hepatitis
27	564.6	89.2	9379	14 AR165051	AR165051 Hepatitis
28	564.6	89.2	9448	14 HPCJ483	D13558 Hepatitis C
29	564.6	89.2	9605	14 HCC238799	AJ38799 Hepatitis
30	564.6	89.2	11076	6 AX036252	AX036252 Sequence
31	564.6	89.2	11076	6 AX036258	AX036258 Sequence
32	564.6	89.2	11076	6 AX036260	AX036260 Sequence
33	564.6	89.2	11076	6 AX036262	AX036262 Sequence
34	564.2	89.1	9598	14 AB049101	AB049101 Hepatitis
35	563	88.9	9033	14 HCC238800	AJ38800 Hepatitis
36	563	88.9	9379	14 AR165048	AR165048 Hepatitis
37	562.6	88.9	9435	14 AB049093	AB049093 Hepatitis
38	561.4	88.7	932	6 E06695	E06695 non A non B
39	561.4	88.7	932	6 I49749	I49749 Sequence 15
40	561.4	88.7	932	6 I63387	I63387 Sequence 15
41	561.4	88.7	932	6 S46012	S46012 polypeptide
42	561.4	88.7	1539	6 AR027786	AR027786 Sequence
43	561.4	88.7	1625	14 HPCNSISPG	M74810 Hepatitis C
44	561.4	88.7	1863	6 AR027783	AR027783 Sequence
45	561.4	88.7	3296	14 AB008441	AB008441 Hepatitis

ALIGNMENTS

RESULT 1
LOCUS A48669 633 bp DNA linear PAT 07-MAR-1997
DEFINITION Sequence 7 from Patent WO9604385.
ACCESSION A48669
VERSION A48669.1 GI:2302382
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 633)
AUTHORS Maertens G., Bosman F., De M.G. and Buysse M.
TITLE PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
JOURNAL Patent: WO 9604385-A 7 15-FEB-1996;

INNOGENETICS NV (BE)
 Other publication CA 2172273 960215
 Other publication AU 3382495 960304.
 FEATURES
 source
 1. 633
 /organism="unidentified"
 /db_xref="taxon:32644"
 1. 630
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAA03130.1"
 /db_xref="gi:2302383"
 /translation="MLGKVIDPLNCFADIMKGYPLVGNAPLGGARALAHGVRLEDG
 VYATGNLPGCSFSEFTLLSLTIPASAYEVRVSGMYHRTNCCSSSTYEAADM
 IMHRTGCTPCVRENSSRCWALPTTLARNA SVPTTTRRHVDLLVGAAPCSAMIV
 GDLCGSLVLSOLFETISPRRHETVODNCSTYPGHITGHRAWMDMMNW"
 1. 627
 /product="unnamed"
 BASE COUNT 111 a 192 c 174 g 156 t
 ORIGIN
 Query Match 100.0%; Score 633; DB 6; Length 633;
 Best Local Similarity 100.0%; Pred. No. 6.4e-144;
 Matches 633; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGTTGGTAAGGTATCATACCTTACGTGGGCTTGGCGACTCATGGGTACATT 60
 1 ATGTTGGTAAGGTATCATACCTTACGTGGGCTTGGCGACTCATGGGTACATT 60
 61 CCGCTGTGGGGCCCCCTAGGGGGTGTGCCAGAGCCCTGGCGCATGGCGTGGGGTT 120
 61 CCGCTGTGGGGCCCCCTAGGGGGTGTGCCAGAGCCCTGGCGCATGGCGTGGGGTT 120
 121 CTGGAAGAGCGGGTGAACATATGCAACAGGAAATTGCTGTGCTCTTCTATCTTC 180
 121 CTGGAAGAGCGGGTGAACATATGCAACAGGAAATTGCTGTGCTCTTCTATCTTC 180
 121 CTGGAAGAGCGGGTGAACATATGCAACAGGAAATTGCTGTGCTCTTCTATCTTC 180
 181 CTCTTGCTTTACTGCTGCTGTGACCATTCAGCTTCGGCTTATAGTGGCAACGTG 240
 181 CTCTTGCTTTACTGCTGCTGTGACCATTCAGCTTCGGCTTATAGTGGCAACGTG 240
 181 CTCTTGCTTTACTGCTGCTGTGACCATTCAGCTTCGGCTTATAGTGGCAACGTG 240
 241 TCCGGGATGTACCATGTCAAGAACGACTGCTCAACTCAAGCATTTGTATGAGGACGG 300
 241 TCCGGGATGTACCATGTCAAGAACGACTGCTCAACTCAAGCATTTGTATGAGGACGG 300
 301 GACATGATCATGCACACCCCGGGGTGCTGCCCTGCGTTGCGGAGAACAACTTCCCGC 360
 301 GACATGATCATGCACACCCCGGGGTGCTGCCCTGCGTTGCGGAGAACAACTTCCCGC 360
 361 TGTGGGTAGCGCTCAACCCCGAGCTGCGAGTGAAGAACGCCAGCTCCCACTACGACA 420
 361 TGTGGGTAGCGCTCAACCCCGAGCTGCGAGTGAAGAACGCCAGCTCCCACTACGACA 420
 421 ATACGACGCCACGTCGATTTGCTGTTGGGGGGCTGTTTCTGTTCCGCTATGACGTG 480
 421 ATACGACGCCACGTCGATTTGCTGTTGGGGGGCTGTTTCTGTTCCGCTATGACGTG 480
 421 ATACGACGCCACGTCGATTTGCTGTTGGGGGGCTGTTTCTGTTCCGCTATGACGTG 480
 481 GGGGATCTTGGGGATCTGCTTCTCTGCTCCAGCTGTTCAACATCTCCGCTCCCGG 540
 481 GGGGATCTTGGGGATCTGCTTCTCTGCTCCAGCTGTTCAACATCTCCGCTCCCGG 540
 541 CATGAGACGGTGCAGACATGCAATTTGCTCAATCTATCCGGCCACATTAACAGGTACCGT 600
 541 CATGAGACGGTGCAGACATGCAATTTGCTCAATCTATCCGGCCACATTAACAGGTACCGT 600
 601 ATGGCTTGGGATATGATGATGAAGTGGTAATAG 633
 601 ATGGCTTGGGATATGATGATGAAGTGGTAATAG 633

DEFINITION Sequence 7 from patent US 6245503.
 ACCESSION AR157326
 VERSION AR157326.1 GI:16218259
 KEYWORDS
 SOURCE
 ORGANISM
 Unknown.
 Unclassified.
 1 (bases 1 to 633)
 REFERENCES
 Maertens/G., Bosman, F., De-Martynoff, G. and Buyse, M.-A.
 Purified hepatitis C virus envelope proteins for diagnostic and
 therapeutic use.
 JOURNAL
 Patent US 6245503-A 7, 12 JUN 2001
 FEATURES
 source
 1. 633
 /organism="unknown"
 BASE COUNT 111 a 192 c 174 g 156 t
 ORIGIN
 Query Match 100.0%; Score 633; DB 6; Length 633;
 Best Local Similarity 100.0%; Pred. No. 6.4e-144;
 Matches 633; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGTTGGTAAGGTATCATACCTTACGTGGGCTTGGCGACTCATGGGTACATT 60
 1 ATGTTGGTAAGGTATCATACCTTACGTGGGCTTGGCGACTCATGGGTACATT 60
 61 CCGCTGTGGGGCCCCCTAGGGGGTGTGCCAGAGCCCTGGCGCATGGCGTGGGGTT 120
 61 CCGCTGTGGGGCCCCCTAGGGGGTGTGCCAGAGCCCTGGCGCATGGCGTGGGGTT 120
 121 CTGGAAGAGCGGGTGAACATATGCAACAGGAAATTGCTGTGCTCTTCTATCTTC 180
 121 CTGGAAGAGCGGGTGAACATATGCAACAGGAAATTGCTGTGCTCTTCTATCTTC 180
 121 CTGGAAGAGCGGGTGAACATATGCAACAGGAAATTGCTGTGCTCTTCTATCTTC 180
 181 CTCTTGCTTTACTGCTGCTGTGACCATTCAGCTTCGGCTTATAGTGGCAACGTG 240
 181 CTCTTGCTTTACTGCTGCTGTGACCATTCAGCTTCGGCTTATAGTGGCAACGTG 240
 181 CTCTTGCTTTACTGCTGCTGTGACCATTCAGCTTCGGCTTATAGTGGCAACGTG 240
 241 TCCGGGATGTACCATGTCAAGAACGACTGCTCAACTCAAGCATTTGTATGAGGACGG 300
 241 TCCGGGATGTACCATGTCAAGAACGACTGCTCAACTCAAGCATTTGTATGAGGACGG 300
 301 GACATGATCATGCACACCCCGGGGTGCTGCCCTGCGTTGCGGAGAACAACTTCCCGC 360
 301 GACATGATCATGCACACCCCGGGGTGCTGCCCTGCGTTGCGGAGAACAACTTCCCGC 360
 361 TGTGGGTAGCGCTCAACCCCGAGCTGCGAGTGAAGAACGCCAGCTCCCACTACGACA 420
 361 TGTGGGTAGCGCTCAACCCCGAGCTGCGAGTGAAGAACGCCAGCTCCCACTACGACA 420
 421 ATACGACGCCACGTCGATTTGCTGTTGGGGGGCTGTTTCTGTTCCGCTATGACGTG 480
 421 ATACGACGCCACGTCGATTTGCTGTTGGGGGGCTGTTTCTGTTCCGCTATGACGTG 480
 421 ATACGACGCCACGTCGATTTGCTGTTGGGGGGCTGTTTCTGTTCCGCTATGACGTG 480
 481 GGGGATCTTGGGGATCTGCTTCTCTGCTCCAGCTGTTCAACATCTCCGCTCCCGG 540
 481 GGGGATCTTGGGGATCTGCTTCTCTGCTCCAGCTGTTCAACATCTCCGCTCCCGG 540
 541 CATGAGACGGTGCAGACATGCAATTTGCTCAATCTATCCGGCCACATTAACAGGTACCGT 600
 541 CATGAGACGGTGCAGACATGCAATTTGCTCAATCTATCCGGCCACATTAACAGGTACCGT 600
 601 ATGGCTTGGGATATGATGATGAAGTGGTAATAG 633
 601 ATGGCTTGGGATATGATGATGAAGTGGTAATAG 633

RESULT 2
 LOCUS AR157326 633 bp DNA linear PAT 17-OCT-2001

RESULT 3
 LOCUS AX452756 633 bp DNA linear PAT 06-JUL-2002
 DEFINITION Sequence 7 from Patent EP1201305
 ACCESSION AX452756 GI:24712441

KEYWORDS
SOURCE Hepatitis C virus.
ORGANISM Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.

REFERENCE
1 Maertens,G., Bosman,F., de Martynoff,G. and Buysse,M.A.
TITLE Recombinant vectors for producing hcv envelope proteins.
JOURNAL Patent: EP 1214315-A 7-05-JUN-2002
Innogenetics N.V. (BE)

FEATURES
source
1. 633
/organism="Hepatitis C virus"
/db_xref="taxon:11103"
1. .630
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD38068.1"
/db_xref="GI:21712442"
/translation="MLGRVLDLTCGFADLMGYIDLVGAPLGGAARALAHGVRLDGVNATGNDLPGCSFSLFLALISCLTIPASAEVRNVSGMYHTNDSSNSIYEADIMHTPGCVPCVRENNSSRCWVALPTPLAARNASVPTTIRRHVDLVGAAPFCSAMYGDLCGVFLVSQLFTISPRRHETVQDCNSIYPGHITGHRAMDMMMNM"

mat_peptide
1. 627
/product="unnamed"

BASE COUNT 111 a 192 c 174 g 156 t
ORIGIN

Query Match 100.0%; Score 633; DB 6; Length 633;
Best Local Similarity 100.0%; Pred. No. 6.4e-144;
Matches 633; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGGGTAAGTCAATCATACCTTACGTGCGGCTTCGCCACCTCATGGGATATT 60
DB 1 ATGTTGGGTAAGTCAATCATACCTTACGTGCGGCTTCGCCACCTCATGGGATATT 60
QY 61 CCGCTGCTGGGGCCCCCTAGGGGGTGTGCCAGAGCCCTGGCGATGGCGTGGGTT 120
DB 61 CCGCTGCTGGGGCCCCCTAGGGGGTGTGCCAGAGCCCTGGCGATGGCGTGGGTT 120
QY 121 CTGGAAGAGGGGGTGAACATGCAACAGGAATTGCGTGGTTCCTTCTATCTTC 180
DB 121 CTGGAAGAGGGGGTGAACATGCAACAGGAATTGCGTGGTTCCTTCTATCTTC 180
QY 181 CTTTGGCTTTACTGCTGCTGACCAATTCAGCTTCGGTTATGAGTGGCAACGTG 240
DB 181 CTTTGGCTTTACTGCTGCTGACCAATTCAGCTTCGGTTATGAGTGGCAACGTG 240
QY 241 TCCGGGATGTACCATGTCAAGACGACTGCTCCAACTCAAGCATTTGTATGAGGACGG 300
DB 241 TCCGGGATGTACCATGTCAAGACGACTGCTCCAACTCAAGCATTTGTATGAGGACGG 300
QY 301 GACATGATATGACACACCCCGGGTGGCCCTGGTTCGGAGAGAACATCTTCCCGC 360
DB 301 GACATGATATGACACACCCCGGGTGGCCCTGGTTCGGAGAGAACATCTTCCCGC 360
QY 361 TCGTGGGTAGCCCTACCCCGAGCTCGAGCTAGGAAGCCAGCGCTCCCACTAGAGAA 420
DB 361 TCGTGGGTAGCCCTACCCCGAGCTCGAGCTAGGAAGCCAGCGCTCCCACTAGAGAA 420
QY 421 ATACGACGCCACGTGATTTGCTGTTGGGGGGGCTGTTTCTGTTCCGCTATGACGTG 480
DB 421 ATACGACGCCACGTGATTTGCTGTTGGGGGGGCTGTTTCTGTTCCGCTATGACGTG 480
QY 481 GGGGATCTTGGGATCTGTTCTGCTGCTCCAGCTTTCACCATCTCGGCTCGCGGG 540
DB 481 GGGGATCTTGGGATCTGTTCTGCTGCTCCAGCTTTCACCATCTCGGCTCGCGGG 540
QY 541 CATGAGACGGTCCAGAGCTCAATTCATATCCCGGACATTAACAGGTCAACGTT 600
DB 541 CATGAGACGGTCCAGAGCTCAATTCATATCCCGGACATTAACAGGTCAACGTT 600
QY 601 ATGGCTTGGGATATGATGATGAATGCTGTAATAG 633

DB 601 ATGGCTTGGGATATGATGATGAATGCTGTAATAG 633

RESULT 4
LOCUS A48675
DEFINITION Sequence 13 from Patent WO9604385.
ACCESSION A48675
VERSION A48675.1 GI:2302388
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 636)
Maertens,G., Bosman,F., De,M.G. and Buysse,M.
TITLE PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
JOURNAL Patent: WO 9604385-A 13 15-FEB-1996;
INNOCENTICS NV (BE)
COMMENT Other publication CA 2172273 960215
Other publication AU 3382495 960304.
FEATURES
source
1. 636
/organism="unidentified"
/db_xref="taxon:32644"
1. .633
/note="unnamed protein product"
/codon_start=1
/protein_id="CAA03133.1"
/db_xref="GI:2302389"
/translation="MLGRAIDLTTCGFADLMGYIDLVGAPLGGAARALAHGVRLDGVNATGNDLPGCSFSLFLALISCLTIPASAEVRNVSGMYHTNDSSNSIYEADIMHTPGCVPCVRENNSSRCWVALPTPLAARNASVPTTIRRHVDLVGAAPFCSAMYGDLCGVFLVSQLFTISPRRHETVQDCNSIYPGHITGHRAMDMMMNM"

mat_peptide
1. 630
/product="unnamed"

BASE COUNT 111 a 197 c 175 g 153 t
ORIGIN

Query Match 96.8%; Score 612.8; DB 6; Length 636;
Best Local Similarity 98.1%; Pred. No. 5.3e-139;
Matches 620; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 ATGTTGGGTAAGTCAATCATACCTTACGTGCGGCTTCGCCACCTCATGGGATATT 60
DB 1 ATGTTGGGTAAGTCAATCATACCTTACGTGCGGCTTCGCCACCTCATGGGATATT 60
QY 61 CCGCTGCTGGGGCCCCCTAGGGGGTGTGCCAGAGCCCTGGCGATGGCGTGGGTT 120
DB 61 CCGCTGCTGGGGCCCCCTAGGGGGTGTGCCAGAGCCCTGGCGATGGCGTGGGTT 120
QY 121 CTGGAAGAGGGGGTGAACATGCAACAGGAATTGCGTGGTTCCTTCTATCTTC 180
DB 121 CTGGAAGAGGGGGTGAACATGCAACAGGAATTGCGTGGTTCCTTCTATCTTC 180
QY 181 CTTTGGCTTTACTGCTGCTGACCAATTCAGCTTCGGTTATGAGTGGCAACGTG 240
DB 181 CTTTGGCTTTACTGCTGCTGACCAATTCAGCTTCGGTTATGAGTGGCAACGTG 240
QY 241 TCCGGGATGTACCATGTCAAGACGACTGCTCCAACTCAAGCATTTGTATGAGGACGG 300
DB 241 TCCGGGATGTACCATGTCAAGACGACTGCTCCAACTCAAGCATTTGTATGAGGACGG 300
QY 301 GACATGATATGACACACCCCGGGTGGCCCTGGTTCGGAGAGAACATCTTCCCGC 360
DB 301 GACATGATATGACACACCCCGGGTGGCCCTGGTTCGGAGAGAACATCTTCCCGC 360
QY 361 TCGTGGGTAGCCCTACCCCGAGCTCGAGCTAGGAAGCCAGCGCTCCCACTAGAGAA 420
DB 361 TCGTGGGTAGCCCTACCCCGAGCTCGAGCTAGGAAGCCAGCGCTCCCACTAGAGAA 420
QY 421 ATACGACGCCACGTGATTTGCTGTTGGGGGGGCTGTTTCTGTTCCGCTATGACGTG 480

Db 301 GACATGATCATGACACACCCCGGGTGGTCCCTGGCTTGGGAGAACACTCTTCCCGC 360
 QY 361 TGGTGGGTAGCGGTCAACCCCGACGCTGCGACGTAGGAACCCAGCGTCCCTACAGACA 420
 Db 361 TGGTGGGTAGCGGTCAACCCCGACGCTGCGACGTAGGAACCCAGCGTCCCTACAGACA 420
 QY 421 ATACGACGCGACGCTGATTTGCTGCTGGGGGGGCTGCTTCTGTTCCGCTATGACGTG 480
 Db 421 ATACGACGCGACGCTGATTTGCTGCTGGGGGGGCTGCTTCTGTTCCGCTATGACGTG 480
 QY 481 GGGGATCTCGGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
 Db 481 GGGGATCTCGGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
 QY 541 CATGAGACGCTGCGACGCTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
 Db 541 CATGAGACGCTGCGACGCTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
 QY 601 ATGGCTTGGGATATGATGATGAAGTGGTAATA 632
 Db 601 ATGGCTTGGGATATGATGATGAAGTGGTAATA 632

RESULT 7
 LOCUS A48667 795 bp DNA linear PAT 07-MAR-1997
 DEFINITION Sequence 5 from Patent WO9604385.
 ACCESSION A48667
 VERSION A48667.1 GI:2302380
 KEYWORDS
 SOURCE
 ORGANISM
 unclassified.
 unclassified.
 1 (bases 1 to 795)
 Maertens,G., Bosman,F., De,M.G. and Buyse,M.
 PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND
 THERAPEUTIC USE

REFERENCE
 AUTHORS Maertens,G., Bosman,F., De,M.G. and Buyse,M.
 TITLE PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND
 JOURNAL THERAPEUTIC USE
 COMMENT Patent: WO 9604385-A-5 15-FEB-1996.
 INNOCENTICS NV (BE)
 Other publication CA 2172273 960215
 Other publication AU 3382495 960304.
 FEATURES
 source
 1..795
 Location/Qualifiers
 /organism="unclassified"
 /db_xref="taxon:32644"
 1..792
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAA03129.1"
 /db_xref="GI:2302381"
 /translation="MGKYIDPLTCGFADLVGIPLVGAPLGGARALAHGVRLIEDG
 VNYATGNLPGCSFISFLALISCLTVPASAYVRNVSQVHTNDCSNSIYEAADM
 IMHTPGCVPCVRENNSSRCQWALPTPLAARAYSVPTTTRRVDLLNCAACSAAMV
 GDLCGSYFLVSLFTISPRRHETVDDCKNSITPGHITGRAMDMAMNSPTTALVVS
 QLRIPQAVNDVAGAHWGLAGLAIVSYVGMARVLLVMLFAP"
 1..789
 /product="unnamed"
 /organism="unknown"

BASE COUNT 130 a 240 c 231 g 194 t
 ORIGIN
 Query Match 96.2%; Score 608.8; DB 6; Length 795;
 Best Local Similarity 98.1%; Pred. No. 5.1e-138;
 Matches 616; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 ATGTTGGTAAAGTCAATCATACCTTACGTGGGCTTGGCCGACCTCATAGGGTACATT 60
 Db 1 ATGTTGGTAAAGTCAATCATACCTTACGTGGGCTTGGCCGACCTCATAGGGTACATT 60
 QY 61 CCGCTGCTGGGCGCCCGCTAGGGGGTGTGCCAGAGCCCTGGCGCATGGCGTCCGGATT 120
 Db 61 CCGCTGCTGGGCGCCCGCTAGGGGGTGTGCCAGAGCCCTGGCGCATGGCGTCCGGATT 120
 QY 121 CTGGAAGACGGCGTGAACATATGCAACAGGAAATTGCTGCTGCTTCTCTATCTTC 180

Db 121 CTGGAAGACGGCGTGAACATATGCAACAGGAAATTGCTGCTGCTTCTCTATCTTC 180
 QY 181 CTCTGGCTTTACTGTCGCTGTCGACCAATTCACGCTTCGCGCTATAGAGGGCCACGTTG 240
 Db 181 CTCTGGCTTTACTGTCGCTGTCGACCAATTCACGCTTCGCGCTATAGAGGGCCACGTTG 240
 QY 241 TCCGGATGATACCATGTCACAGACGACTGCTCCACACTCAAGCATTTGTATAGAGCAGG 300
 Db 241 TCCGGATGATACCATGTCACAGACGACTGCTCCACACTCAAGCATTTGTATAGAGCAGG 300
 QY 301 GACATATCATGACACACCCCGGGTGGTCCCTGGCTTGGGAGAACACTTCTCCCGC 360
 Db 301 GACATATCATGACACACCCCGGGTGGTCCCTGGCTTGGGAGAACACTTCTCCCGC 360
 QY 361 TGGTGGGTAGCGGTCAACCCCGACGCTGCGACGTAGGAACCCAGCGTCCCTACAGACA 420
 Db 361 TGGTGGGTAGCGGTCAACCCCGACGCTGCGACGTAGGAACCCAGCGTCCCTACAGACA 420
 QY 421 ATACGACGCGACGCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
 Db 421 ATACGACGCGACGCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
 QY 481 GGGGATCTCGGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
 Db 481 GGGGATCTCGGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
 QY 541 CATGAGACGCTGCGACGCTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
 Db 541 CATGAGACGCTGCGACGCTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
 QY 601 ATGGCTTGGGATATGATGATGAAGTGGT 628
 Db 601 ATGGCTTGGGATATGATGATGAAGTGGT 628

RESULT 8
 LOCUS ARI57325 795 bp DNA linear PAT 17-OCT-2001
 DEFINITION Sequence 5 from patent US_6245503-
 ACCESSION ARI57325
 VERSION ARI57325.1 GI:16218258
 KEYWORDS
 SOURCE
 ORGANISM
 Unknown.
 Unknown.
 unclassified.
 1 (bases 1 to 795)
 Maertens,G., Bosman,F., De Martynoff,G. and Buyse,M.-A.
 Purified hepatitis C virus envelope proteins for diagnostic and
 therapeutic use
 JOURNAL Patent: US 6245503-A 5 12-JUN-2001;
 FEATURES
 source
 1..795
 Location/Qualifiers
 /organism="unknown"
 /product="unnamed"
 /organism="unknown"

BASE COUNT 130 a 240 c 231 g 194 t
 ORIGIN
 Query Match 96.2%; Score 608.8; DB 6; Length 795;
 Best Local Similarity 98.1%; Pred. No. 5.1e-138;
 Matches 616; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 ATGTTGGTAAAGTCAATCATACCTTACGTGGGCTTGGCCGACCTCATAGGGTACATT 60
 Db 1 ATGTTGGTAAAGTCAATCATACCTTACGTGGGCTTGGCCGACCTCATAGGGTACATT 60
 QY 61 CCGCTGCTGGGCGCCCGCTAGGGGGTGTGCCAGAGCCCTGGCGCATGGCGTCCGGATT 120
 Db 61 CCGCTGCTGGGCGCCCGCTAGGGGGTGTGCCAGAGCCCTGGCGCATGGCGTCCGGATT 120
 QY 121 CTGGAAGACGGCGTGAACATATGCAACAGGAAATTGCTGCTGCTTCTCTATCTTC 180
 Db 121 CTGGAAGACGGCGTGAACATATGCAACAGGAAATTGCTGCTGCTTCTCTATCTTC 180

Query Match	Best Local Similarity	96.1%	Score 608.8	DB 6	Length 795
Matches 616	Conservative 0	Mismatches 12	Indels 0	Gaps 0	
<p>1 ATGTTGGGTAAGGTCATGCATACCTTACGTCGGGCTTCGCCGACCTCATGGGTCACATT 60</p>					

Db	1	ATGTTGGTAAAGCATGATACCTTACATACGCGCTTGGGAGCTGTTGGGGACTT	60
Qy	61	CCGCTCGTCGCGGCCCCCTAGGGGGTGCCTGCACAGGCTTCGGCAGTGGCGTCCGGTT	120
Db	61	CCGCTCGTCGCGGCCCCCTAGGGGGGCGTGCACAGGCGCTTGGCAGTGGCGTCCGGTT	120
Qy	121	CTGGAAAGACGGCGTGAATCATATGCAACAGGGAATTTGCCGTGTTCTTCTATCTTC	180
Db	121	CTGGAGGAGAGGGGTGAACATATGCAACAGGGAATTTGCCGTGTTCTTCTATCTTC	180
Qy	181	CTCTTGGCTTCTACTGCTGCTGCATTCACGTTCCGCTTATGAGGTGGCAAGCTG	240
Db	181	CTCTTGGCTTCTACTGCTGCTGCATTCACGTTCCGCTTATGAGGTGGCAAGCTG	240
Qy	241	TCCGGGATGTACCATGTATGCAAGCAACTGCTCCAACTAAGCATTTGTATGAGGACAG	300
Db	241	TCCGGGATGTACCATGTATGCAAGCAACTGCTCCAACTAAGCATTTGTATGAGGACAG	300
Qy	301	GACATGATCATGACACACCCCGGGTGCCTGCTCGTTGGGAGAACACTCTTCCGC	360
Db	301	GACATGATCATGACACACCCCGGGTGCCTGCTCGTTGGGAGAACACTCTTCCGC	360
Qy	361	TGCTGGGTAGGCGCTACCCCGCAGCTCCGAGTAAAGGACGAGGTCCCACTACGACA	420
Db	361	TGCTGGGTAGGCGCTACCCCGCAGCTCCGAGTAAAGGACGAGGTCCCACTACGACA	420
Qy	421	ATACAGCGCACAGTGGATTTGCTGCTGGGGGGGCGCTTTCTTTCCGCTATGACGTG	480
Db	421	ATACAGCGCACAGTGGATTTGCTGCTGGGGGGGCGCTTTCTTTCCGCTATGACGTG	480
Qy	481	GGGGATCTCTGGGATCTGTCTCTCTGCTCCAGCTGTTACCATCTCGGCTCGCGG	540
Db	481	GGGGATCTCTGGGATCTGTCTCTCTGCTCCAGCTGTTACCATCTCGGCTCGCGG	540
Qy	541	CATGAGACGCTGCAAGACTGCAATTTGCTCAATCTATCCCGGCAATTAAGGTACCGT	600
Db	541	CATGAGACGCTGCAAGACTGCAATTTGCTCAATCTATCCCGGCAATTAAGGTACCGT	600
Qy	601	ATGCGTTGGATATGATGATGAACGTG	628
Db	601	ATGCGTTGGATATGATGATGAACGTG	628

VNYATGNLPCCSFIFLALLSCLTVPASATVENVSCMTHTVNDSCNSIYEADMD
IMHTPCVPCVRENNSRCVVALTPPLAARNASVPTT11RRHVDLLVGAACASMY
GDLCSVEFVQLTISPRHEVODNCISIPGHITGRHMDMMNNSPTTALVVS
OLLRIPOAVDMVAGAHGVLAGLAYSVWGVVLYVLMFLFAGDITRVSAGAA
SDTRGLVSLFSPGSOKIOIYVNTGSMHNRALNCNSIQGFPAALFYKRFNSG
CPERLASCRSIDKPAQGMPLATTPRNSDDQPYCMTWHPRRCGIVPAVSQVPCYFC
TPSPVVGTTDRFVPTYNMGANDSDVILLNTRPPRGWFCCTMMNGGFKTCGGP
PCNIGAGNNMTLTCPTDCRKHPEATVARGSGPWLTPRCMVHYPRILMHYPCVNTF
IFKVMYVGVGHEHREACNMTRGERCELDRESELSPSLSTEOMLILPCSEFTLLP
ALSTGLIHONHIVDOYLYGVSAVSLVITWEYVELLFLLDARICACIMMLLI
NOAEALENLYVLMNAAYAGHGTSLFLVFFCAANYIKGRLYPGAAYFYGWPLILL
LLALPPRAVA
1..2076
mat_peptide /product="unnamed"

BASE COUNT 366 a 634 c 600 g 482 t
ORIGIN

Query Match 95.7%; Score 605.8; DB 6; Length 2082;
Best Local Similarity 98.1%; Pred. No. 2.9e-137;

Matches 613; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 4 TTGGGTAAAGTCATCATACCTTACGTGGGGCTTGCCGACCTCATGGGTCATTCCG 63
DB 4 TTGGGTAAAGTCATCATACCTTACGTGGGGCTTGCCGACCTCATGGGTCATTCCG 63
QY 64 CTCGTGGGGCCCCCTAGGGGGTCTGCCAGAGCCCTGGCCATGGCGTTCGGTTCG 123
DB 64 CTCGTGGGGCCCCCTAGGGGGTCTGCCAGAGCCCTGGCCATGGCGTTCGGTTCG 123
QY 124 GAAGACGGCGTCACTATGCAACAGGGAATTTCCGTGCTTCTCTATCTTCCTC 183
DB 124 GAAGACGGCGTCACTATGCAACAGGGAATTTCCGTGCTTCTCTATCTTCCTC 183
QY 184 TTGGCTTACTGTCGTCTGACCATTCAGCTTCGGCTATAGAGGCGCCACGTCGC 243
DB 184 TTGGCTTACTGTCGTCTGACCATTCAGCTTCGGCTATAGAGGCGCCACGTCGC 243
QY 244 GGGATATACATGTCACGAAGCACTGCTCCAACTCAAGCATTTGTATAGAGCGAGC 303
DB 244 GGGATATACATGTCACGAAGCACTGCTCCAACTCAAGCATTTGTATAGAGCGAGC 303
QY 304 ATGATATGACACACCCCGGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363
DB 304 ATGATATGACACACCCCGGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363
QY 364 TTGGGTAGCGTCAACCCCAAGCTGAGCTAGGAAGCCAGCGTCCCACTAGACAATA 423
DB 364 TTGGGTAGCGTCAACCCCAAGCTGAGCTAGGAAGCCAGCGTCCCACTAGACAATA 423
QY 424 CGACGCCACGTCGATTTGCTGTTGGGGGGGCTGTTTCTGTTCCGCTATGTCAGTGGG 483
DB 424 CGACGCCACGTCGATTTGCTGTTGGGGGGGCTGTTTCTGTTCCGCTATGTCAGTGGG 483
QY 484 GATCTGTGGGATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 543
DB 484 GATCTGTGGGATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 543
QY 544 GAGACGGTGAAGAGTCGCAATTCCTCAATCTATCCGGCCACATAAGGTCACCGTATG 603
DB 544 GAGACGGTGAAGAGTCGCAATTCCTCAATCTATCCGGCCACATAAGGTCACCGTATG 603
QY 604 GCTTGGGATATGATGATGAACCTGCT 628
DB 604 GCTTGGGATATGATGATGAACCTGCT 628

RESULT 11
LOCUS ARI57350 2082 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 47 from patent US 6245503.
ACCESSION ARI57350
VERSION ARI57350.1 GI:16218284
KEYWORDS

SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2082)
AUTHORS Maertens, G., Bosman, F., De Martynoff, G. and Buyse, M.-A.
TITLE Purified hepatitis C virus envelope proteins for diagnostic and therapeutic use
JOURNAL Patent: US 6245503-A 47 12-JUN-2001;
FEATURES location/Qualifiers
source 1..2082
BASE COUNT 366 a 634 c 600 g 482 t
ORIGIN

Query Match 95.7%; Score 605.8; DB 6; Length 2082;
Best Local Similarity 98.1%; Pred. No. 2.9e-137;

Matches 613; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 4 TTGGGTAAAGTCATCATACCTTACGTGGGGCTTGCCGACCTCATGGGTCATTCCG 63
DB 4 TTGGGTAAAGTCATCATACCTTACGTGGGGCTTGCCGACCTCATGGGTCATTCCG 63
QY 64 CTCGTGGGGCCCCCTAGGGGGTCTGCCAGAGCCCTGGCCATGGCGTTCGGTTCG 123
DB 64 CTCGTGGGGCCCCCTAGGGGGTCTGCCAGAGCCCTGGCCATGGCGTTCGGTTCG 123
QY 124 GAAGACGGCGTCACTATGCAACAGGGAATTTCCGTGCTTCTCTATCTTCCTC 183
DB 124 GAAGACGGCGTCACTATGCAACAGGGAATTTCCGTGCTTCTCTATCTTCCTC 183
QY 184 TTGGCTTACTGTCGTCTGACCATTCAGCTTCGGCTATAGAGGCGCCACGTCGC 243
DB 184 TTGGCTTACTGTCGTCTGACCATTCAGCTTCGGCTATAGAGGCGCCACGTCGC 243
QY 244 GGGATATACATGTCACGAAGCACTGCTCCAACTCAAGCATTTGTATAGAGCGAGC 303
DB 244 GGGATATACATGTCACGAAGCACTGCTCCAACTCAAGCATTTGTATAGAGCGAGC 303
QY 304 ATGATATGACACACCCCGGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363
DB 304 ATGATATGACACACCCCGGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363
QY 364 TTGGGTAGCGTCAACCCCAAGCTGAGCTAGGAAGCCAGCGTCCCACTAGACAATA 423
DB 364 TTGGGTAGCGTCAACCCCAAGCTGAGCTAGGAAGCCAGCGTCCCACTAGACAATA 423
QY 424 CGACGCCACGTCGATTTGCTGTTGGGGGGGCTGTTTCTGTTCCGCTATGTCAGTGGG 483
DB 424 CGACGCCACGTCGATTTGCTGTTGGGGGGGCTGTTTCTGTTCCGCTATGTCAGTGGG 483
QY 484 GATCTGTGGGATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 543
DB 484 GATCTGTGGGATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 543
QY 544 GAGACGGTGAAGAGTCGCAATTCCTCAATCTATCCGGCCACATAAGGTCACCGTATG 603
DB 544 GAGACGGTGAAGAGTCGCAATTCCTCAATCTATCCGGCCACATAAGGTCACCGTATG 603
QY 604 GCTTGGGATATGATGATGAACCTGCT 628
DB 604 GCTTGGGATATGATGATGAACCTGCT 628

RESULT 12
LOCUS AX452796 2082 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 47 from Patent EP1211315.
ACCESSION AX452796
VERSION AX452796.1 GI:21712481
KEYWORDS
SOURCE Hepatitis C virus.
ORGANISM Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;


```

|||||
Db 535 TTGGCTTGGCTGCTCTGCTGACCGCTTCACGCTTCGCTTATGAAAGTGGCAAGCTGTCC 594
OY 244 GGGATGTACCATGTGACGAAGCACTGCTCCAACTCAAGCAATTGTATGAGAGCGAGAC 303
Db 595 GGGATGTACCATGTGACGAAGCACTGCTCCAACTCAAGCAATTGTATGAGAGCGAGAC 654
OY 304 ATGATCATGTGACACACCCCGGGGTGGTCCCTGCTTCGGGAGAAACAATCTTCCCGTGC 363
Db 655 ATGATCATGTGACACACCCCGGGGTGGTCCCTGCTTCGGGAGAAACAATCTTCCCGTGC 714
OY 364 TTGGTACGCTCAACCCCGAGCTGCGAGCTAGAGAACGCCAGCGTCCCACTACACATA 423
Db 715 TTGGTACGCTCAACCCCGAGCTGCGAGCTAGAGAACGCCAGCGTCCCACTACACATA 774
OY 424 CGAGCCACGCTCAATTTGCTGCTGGGGGGGCTGCTTCTGTGCTCCGCTATGATGAGGG 483
Db 775 CGAGCCACGCTCAATTTGCTGCTGGGGGGGCTGCTTCTGTGCTCCGCTATGATGAGGG 834
OY 484 GATCTCTGGGATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 543
Db 835 GACCTCTGGGATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 894
OY 544 GAGACGGTGCAGAGCTCAATTTGCTGCTGGGGGGGCTGCTTCTGTGCTCCGCTATGATG 603
Db 895 GAGACGGTGCAGAGCTCAATTTGCTGCTGGGGGGGCTGCTTCTGTGCTCCGCTATGATG 954
OY 604 GCTTGGGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 628
Db 955 GCTTGGGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 979

RESULT 14
LOCUS AR157351 2433 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 49 from patent US 6245503.
ACCESSION AR157351
VERSION AR157351.1 GI:16218285
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 2433)
Maertens,G., Bosman,F., De Martynoff,G. and Buyse,M.-A.
Purified hepatitis C virus envelope proteins for diagnostic and
therapeutic use
Patent: US 6245503-A 49 12-JUN-2001;
FEATURES
source location/Qualifiers
1..2433
BASE COUNT 434 a 745 c 714 g 540 t
ORIGIN
Query Match 95.7%; Score 605.8; DB 6; Length 2433;
Best Local Similarity 98.1%; Pred. No.2.9e-137;
Matches 613; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
OY 4 TTGGTGAAGTCAATCATGATCCTTACCTGAGGCTTCCGACCTCATGGGATACATCCG 63
Db 355 TTGGTGAAGTCAATCATGATCCTTACCTGAGGCTTCCGACCTCATGGGATACATCCG 414
OY 64 CTCGTGCGCGCCCGCCCTAGGGGGTGTGCTGCAAGAGCCCTGCGCATGGCGTCCGGTCTG 123
Db 415 CTCGTGCGCGCCCGCCCTAGGGGGTGTGCTGCAAGAGCCCTGCGCATGGCGTCCGGTCTG 474
OY 124 GAGAGCGCGCTGAACATTTGCAACAGGAATTTGCTGCTGCTTCTCTATCTTCTCTC 183
Db 475 GAGAGCGCGCTGAACATTTGCAACAGGAATTTGCTGCTGCTTCTCTATCTTCTCTC 534
OY 184 TTGGCTTACTGCTGCTGCTGACCATTCAGCTTCCGCTATGAGAGCGCAAGCTGTCC 243
Db 535 TTGGCTTACTGCTGCTGCTGACCATTCAGCTTCCGCTATGAGAGCGCAAGCTGTCC 594

```

```

OY 244 GCGATGTACCATGTGACGAAGCACTGCTCCAACTCAAGCAATTGTATGAGAGCGAGAC 303
Db 595 GGGATGTACCATGTGACGAAGCACTGCTCCAACTCAAGCAATTGTATGAGAGCGAGAC 654
OY 304 ATGATCATGTGACACACCCCGGGGTGGTCCCTGCTTCGGGAGAAACAATCTTCCCGTGC 363
Db 655 ATGATCATGTGACACACCCCGGGGTGGTCCCTGCTTCGGGAGAAACAATCTTCCCGTGC 714
OY 364 TTGGTACGCTCAACCCCGAGCTGCGAGCTAGAGAACGCCAGCGTCCCACTACACATA 423
Db 715 TTGGTACGCTCAACCCCGAGCTGCGAGCTAGAGAACGCCAGCGTCCCACTACACATA 774
OY 424 CGAGCCACGCTCAATTTGCTGCTGGGGGGGCTGCTTCTGTGCTCCGCTATGATGAGGG 483
Db 775 CGAGCCACGCTCAATTTGCTGCTGGGGGGGCTGCTTCTGTGCTCCGCTATGATGAGGG 834
OY 484 GATCTCTGGGATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 543
Db 835 GACCTCTGGGATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 894
OY 544 GAGACGGTGCAGAGCTCAATTTGCTGCTGGGGGGGCTGCTTCTGTGCTCCGCTATGATG 603
Db 895 GAGACGGTGCAGAGCTCAATTTGCTGCTGGGGGGGCTGCTTCTGTGCTCCGCTATGATG 954
OY 604 GCTTGGGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 628
Db 955 GCTTGGGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 979

RESULT 15
LOCUS AX452798 2433 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 49 from Patent EP1211315.
ACCESSION AX452798
VERSION AX452798.1 GI:21712483
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1
Maertens,G., Bosman,F., de Martynoff,G. and Buyse,M.A.
Recombinant vectors for producing hcv envelope proteins
Patent: EP 1211315-A 49 05-JUN-2002;
FEATURES
source location/Qualifiers
1..2433
CDS
1..2430
/organism="Hepatitis C virus"
/db_xref="taxon:11103"
/note="unnamed protein product"
/codon_start=1
/db_xref="GI:21712483"
/translating="MSTNPKPQKTKRNTNRPPQDKFFGGQIVGGVYLPRGRRL
GVATRTKTSERSDPGRGROPFIPKRRPBRRAAOPEYMPPLGNDMGAGWLLSPRG
SRPMGPTDPRRRSRNLGKIDLTGCFADLVGYIPLVGAIPGABALAHVRLIED
GVNATAGNLPGCSFSEIFLALISCLTPASAVEVRVGSMAHYHTNDSSSTVYAAD
MIMHPGCVPCVRENNSRCWALPTLTLARNAASVPTTIRRHVDLTVAAAFCSAM
VGDLCGVEFLVQLFTISPRHETVODCKSTYPGHITGHRAMDMNMSPTTALV
SOLLRIPAQVDMVAGAHGVLGLAYSMVGNMAVLVLMLEAGVDHRTVSGAA
ASDTRGLVLFSPGSAOKIQLVNTNGSMHINFTALCNDLSLOTGFALFYRHKFNS
GCPERLASCSRIDKFAOGWGPLTYEPNSDORPYCMHAPRPGIVPAISOVCGPYC
FPSPVVGTTDPRFGVPTYNMGANDSDVILNNTPRRGNMGVCTMMNGTGFTKCG
PCNIGAGANNTLTCPTDPCFRKHPATYRCSGPRILPRCAVHPIRYLRMHPCVNF
TTRKRMVGVGVHRPEACNMTRRGRCDLEDRDSELSPLLLSTTEWQILPCSFYTL
PALSTGLHLHONIVDOYLVGVGSAVSVLVTKWEVILLFLFLADARICLMMMLL
IAQEALENLVVLVLAANAAGAHGTLSPVFCAMY IGRVLVPGAAYAFYGVPLLL
LTLALPPRAYA"
1..2427
mat_peptide
BASE COUNT 434 a 745 c 714 g 540 t
ORIGIN

```

Query Match 95.7%; Score 605.8; DB 6; Length 2433;
Best Local Similarity 98.1%; Pred. No. 2.9e-137;
Matches 613; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```
QY 4 TTGGGTAAGTATGATGATACCTTACGTGGGCTTGGCCGACCTCATGGGTACATTCCG 63
    |||||||
Db 355 TTGGGTAAGTATGATGATACCTTACGTGGGCTTGGCCGACCTCATGGGTACATTCCG 414
    |||||||
QY 64 CTGCTGGGCCCCCCTAGGGGGTGTGCGCAGACCCCTGGCGATGGCGTCCGGTTCTG 123
    |||||||
Db 415 CTGCTGGGCCCCCCTAGGGGGGCTGCCAGGGCCCTGGCGATGGCGTCCGGTTCTG 474
    |||||||
QY 124 GAAGACGGCGTGAATATCAACAGGGAATTGCTGTTGCTCTTCTATCTTCTC 183
    |||||||
Db 475 GAGGACGGCGTGAATATCAACAGGGAATTGCTGTTGCTCTTCTATCTTCTC 534
    |||||||
QY 184 TTGGCTTACTGCTGCTGATGACATTCAGCTTCGCTTATGAGTGGCAACGTGCC 243
    |||||||
Db 535 TTGGCTTACTGCTGCTGATGACATTCAGCTTCGCTTATGAGTGGCAACGTGCC 594
    |||||||
QY 244 GGGATGTACCATGTCAAGACGACTGCTCCAACTCAAGCATTTGTATGAGCAGCGAC 303
    |||||||
Db 595 GGGATGTACCATGTCAAGACGACTGCTCCAACTCAAGCATTTGTATGAGCAGCGAC 654
    |||||||
QY 304 ATGATCATGCACACCCCGGGTGTGCTGCTGCTGCTGCGGAGACACTCTTCCGCTGC 363
    |||||||
Db 655 ATGATCATGCACACCCCGGGTGTGCTGCTGCTGCTGCGGAGACACTCTTCCGCTGC 714
    |||||||
QY 364 TGGGTAGCGCTCACCCCGCAGCTGCGAGTAGGAGCGCGCTCCCACTACGACATA 423
    |||||||
Db 715 TGGGTAGCGCTCACCCCGCAGCTGCGAGTAGGAGCGCGCTCCCACTACGACATA 774
    |||||||
QY 424 CGACGCCACGTGATTTGCTGCTGGGGGCGCTTCTGTTCCGCTATGTACGTGGGG 483
    |||||||
Db 775 CGACGCCACGTGATTTGCTGCTGGGGGCGCTTCTGTTCCGCTATGTACGTGGGG 834
    |||||||
QY 484 GATCTGTGCGATCTGCTCTGCTGCTGCCAGCTGTTACCACTGCGCTGCGGGCAT 543
    |||||||
Db 835 GACCTGTGCGATCTGCTCTGCTGCTGCCAGCTGTTACCACTGCGCTGCGGGCAT 894
    |||||||
QY 544 GAGAGGTCAGGACTGCAATGCTCAATCTATCCGGGCACATACAGTCAACCGTATG 603
    |||||||
Db 895 GAGAGGTCAGGACTGCAATGCTCAATCTATCCGGGCACATACAGTCAACCGTATG 954
    |||||||
QY 604 GCTTGGGATATGATGATGAATGCT 628
    |||||||
Db 955 GCTTGGGATATGATGATGAATGCT 979
    |||||||
```

Search completed: February 18, 2003, 06:39:08
Job time : 2149 secs

OY	244	GGAATGTCACATTGCAGAACAGATGGCTCCAAATCAAGCATTTGTGTAGGAGCGAGC	303
Dd	199	:: glytetyrhnlsvallthrlnaspcysserlmsenlerllevallytgllalaaiaasp	218
OY	304	ATGATCATGCACACCCCCGGGTGGTACCSTGCCGTTCGGGAGAACAATCTTCCTCCGCTGC	363
Dd	219	Mettilemethisthrproglcysval.procystala.lrgtlunspanseserleargcys	258
OY	364	TGGGTAGCGCTCACCCCACGCTGCGAAGTAAGAACGCCAAGCTCCCACTACGACATA	423
Dd	239	 trypalaleutlrprcthllealdlaalargasnalseserval.prolrhtlrhle	258
OY	424	CGAGGCCACGTCGATTTTGCTGCTGGGGGGCGGCTCTTGTCGCGTAGTAACGTGGGG	483
Dd	259	Alrgatghslvsalarplellevalgilyaladlaalaphecysseralametylualgilv	278
OY	484	GATCTCTGCGGATCTGTCTTCCTCCGTCCAGCTGTTCACAACTTCGCTCCCGGCAAT	543
Dd	279	AsplenucysgliserVal.rphleuleval.sercldm.eurpherhrheser.prolgdrlgnls	298
OY	544	GAGACGGTGCAGACATGTCATATGCTCAATATATCCSGGSCACATACAGTCAACCTATG	603
Dd	299	 giuvhrval.glnasrcysasnscysserlleretrolis.leusertcilynhsarimet	318
OY	604	GCTTGGGATATGATGATGAATCG 627	
Dd	319	 Alatrpnaspetmetmedasttp 326	

XX	AA038279	standard; Protein; 733 AA.
XX	AA038279;	
XX	21-OCT-1993	(first entry)
XX	<u>NANB hepatitis virus</u>	HC-OM gene polypeptide P-733-2.
XX	Non-A; non-B; virus;	polymerase chain reaction; detection;
XX	sensitive; specific; HCV; NANBH.	
XX	Non-A, non-B hepatitis virus.	
XX	<u>JP05091884-A.</u>	
XX	16-APR-1993.	
XX	10-APR-1991;	91JP-0196175.
XX	12-JUN-1990;	90JP-0153401.
XX	08-NOV-1990;	90JP-0304405.
XX	(NAKA/)	NAKAMURA T.
XX	WPI: 1993-199637/25.	
XX	DR N-PSDB; AA043889.	
XX	Antigen related to non-A and non-B hepatitis virus - comprises	
XX	non-translation region comprising 340-341 mols. of nucleotides,	
XX	non-translation region comprising 1885-2551 mols. of	
XX	nucleotides including region 1,149 and, etc.	
XX	Claim 15; Page 36-39; 73pp; Japanese.	
XX	The sequence is that of NANB hepatitis virus HC-OM gene polypeptide	
XX	P-733-2. It may be used in a system for detecting NANB hepatitis	
XX	This method is highly specific and sensitive, and can detect NANB	
XX	hepatitis virus which could not be detected by conventional methods.	
XX	Sequence 733 AA;	
XX	Sequence 733 AA;	

Score:	1094.00	Matches:	203
Percent Similarity:	99.52%	Conservative:	4
Best Local Similarity:	97.60%	Mismatches:	1
Query Match:	94.72%	Indels:	
DB:	14	Gaps:	0

US-09-899-303A-7 (1-633) x AAR38279 (1-733)

[illegible]


```

FT      Region      192..383
FT      /label= ENV
FT      Region      384...810
FT      /label= NS1
FT      Region      811..1031
FT      /label= NS2-NS3
FT      /note= "NS3 N-terminal"
XX
XX      JP06141870-A.
XX
XX      24-MAY-1994.
XX
XX      12-MAR-1992; 92JP-0088140.
XX
XX      12-MAR-1992; 92JP-0088140.
XX
XX      (SANMA ) SANMA KAGAKU KENKYUSHO CO.
XX      (TORU ) TONEN CORP.
XX      (TORU-1) ZH TOKYO RINSHO IGAKU SOGO KENKYUSHO.
XX
XX      WPI; 1994-205028/25.
XX      N-PSDB; AA064069.
XX
XX      DNA coding a Non-A, non-B hepatitis virus antigen - useful for
XX      detecting HCV within serum
XX
XX      Claim 6-10; Page 15-19; 22pp; Japanese.
XX
XX      Hepatitis C virus #4 and #6 genes were isolated (AA064068-69).
XX      Both genes contain the core, ENV, NS1, NS2 and NS3 regions.
XX      A core region fragment is given in AA064067.
XX
XX      Sequence 1031 AA:

Alignment Scores:
Pred. No.: 3,76e-104 Length: 1031
Score: 1081.00 Matches: 199
Percent Similarity: 98.08% Conservative: 5
Best Local Similarity: 95.67% Mismatches: 4
Query Match: 93.59% Indels: 0
DB: 15 Gaps: 0

US-09-899-303a-7 (1-633) x AAR54067 (1-1031)
XX
XX      4 TTGGGTAGGTCATCGATCCCTTACGTGCGGCTTCGCCGACCTCATGGGTCATATCCG 63
XX      |||||||
XX      119 LeuGlyLysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIlePro 138
XX      |||||||
XX      64 CTGCTGCGCGCCCGCTAGCGGGGTGCTGCCAAGACCCCTGGCGCATGGCGGTTCG 123
XX      |||||||
XX      139 LeuValGlyAlaProLeuGlyGlyValAlaArgAlaLeuAlaHisGlyValAlaArgValAl 158
XX      |||||||
XX      124 GAAGACGGGCTGACATATCAACAGGGAATTCGCTGCTTCCTTCTATCTTCCTCC 183
XX      |||||||
XX      159 GluAspGlyValAsnTyrAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeu 178
XX      |||||||
XX      184 TTGGCTTACTGCTCTGTGACCATTCAGCTTCGCTTATGAGGTGGCGAACGTGTC 243
XX      |||||||
XX      179 LeuAlaLeuLeuSerCysLeuThrIleProAlaSerAlaTyrGlyValAlaArgAsnValSer 198
XX      |||||||
XX      244 GGGAGTCACCATGTCACGACAGCATGCTCCAACTCAAGCATTTGTATAGCGCGGAC 303
XX      |||||||
XX      199 GlyIleTyrHisValThrAsnAspCysSerAsnSerSerIleValTyrGluAlaAlaAsp 218
XX      |||||||
XX      304 ATGATCATGCACACCGCGGGGTGCGCTGCTGGGAGAACACATCCCGGCTGC 363
XX      |||||||
XX      219 MetIleMetHisThrProGlyCysValProCysValArgGlyAsnSerSerArgCys 238
XX      |||||||
XX      364 TGGGTAGGTCATCCGACCGCTGCGAGCTAGGAACGCCAGGCTCCCACTACGACAAAT 423
XX      |||||||
XX      239 TrpValAlaLeuThrProThrLeuAlaAlaArgAsnAlaSerValProThrThrAlaIle 258
XX      |||||||
XX      424 CGACGCCACGTCGATTTGCTCGTTGGGCGCGCTCTTCTGTCTCGCTATGTACGTGGG 483

```

```

DB      259 ArgArgHisValAlaAspLeuLeuValGlyAlaAlaAlaPheCysSerAlaMetTyrValGly 278
XX      |||||||
XX      484 GATCTGTGCGGATTCGTCTTCCTGCTCCAGCTGTTACCATGTCGTCGCGGAT 543
XX      |||||||
XX      279 AspLeuCysGlySerValPheLeuValSerGlnLeuPheThrPheSerProArgArgHis 298
XX      |||||||
XX      544 GAGACGGTCAGACACTGCATTCCTCAATCTATCCCGGCGACATGAACAGTCACCGTATG 603
XX      |||||||
XX      299 GluThrIleGlnAspCysAsnCysSerIleTyrProGlyHisValSerGlyHisArgMet 318
XX      |||||||
XX      604 GCTTGGGATATCATGATGACTCG 627
XX      |||||||
XX      319 AlaTrpAspMetMetMetAsnTrp 326
XX      |||||||

RESULT 8
AAR98362
ID      AAR98362 standard; Protein, 1031 AA.
XX
XX      AAR98362;
XX
XX      22-AUG-1996 (first entry)
XX
XX      5'UTR/CORE/ENV/NS1/NS2/NS3 from HCV (#6).
XX
XX      Hepatitis C virus; HCV; antigen; detection; antibody.
XX
XX      Hepatitis C virus.
XX
XX      Key      Location/Qualifiers
XX      FH      Peptide      1..191
XX      FT      Peptide      /label= Core peptide
XX      FT      Peptide      192..383
XX      FT      Peptide      /label= ENV1
XX      FT      Peptide      384..810
XX      FT      Peptide      /label= NS1/NS2
XX      FT      Peptide      811..1031
XX      FT      Peptide      /label= NS2 and NS3

JP07133291-A.
XX
XX      23-MAY-1995.
XX
XX      18-JUN-1993; 93JP-0147944.
XX
XX      18-JUN-1993; 93JP-0147944.
XX
XX      (TORU ) TONEN CORP.
XX
XX      WPI; 1995-220780/29.
XX      N-PSDB; AAT30387.
XX
XX      Recombinant polypeptide comprising partial NS1 region of hepatitis
XX      non-A non-B viral antigen - used in a method for detecting
XX      antibodies against hepatitis non-A non-B virus.
XX
XX      Disclosure; Page 13-15; 15pp; Japanese.
XX
XX      The sequences given in AAR98361-62 represent the 5'UTR/CORE/ENV/NS1/NS2/
XX      NS3 protein region derived from hepatitis C virus (HCV) isolates #4
XX      and #6 respectively. The proteins partic. contain amino acids 384-
XX      495 of the HCV NS1 antigen. These protein fragments may be used in
XX      the detection of antibodies against HCV.
XX
XX      Sequence 1031 AA:

Alignment Scores:
Pred. No.: 3,76e-104 Length: 1031
Score: 1081.00 Matches: 199
Percent Similarity: 98.08% Conservative: 5
Best Local Similarity: 95.67% Mismatches: 4
Query Match: 93.59% Indels: 0
DB: 16 Gaps: 0

```

US-09-899-303a-7 (1-633) x AAM98362 (1-1031)

QY 4 TTGGGTAAGGTCATGATACCTTACGTCGGGCTTCGCCGACCTCATGGGGTACATTCG 63
 DB 119 LeucllyysValIleapThrlleuthrCysGlyphealaapleumetGlyTyrIlePro 138
 QY 64 CTCGTCGGCCGCCCTAGGGGGTGTCCAGAGCCCTGGCGCATGGCGGTCTGTG 123
 DB 139 LeuValIleValProleuGlyGlyValIleAlaArgIleLeuAlaHISGlyValArgValVal 158
 QY 124 GAAGACGGCGGTGACATTCGACAGGGAATTTGGCTGTCTTCTTCTATCTTCCTC 183
 DB 159 GluaspIlyValasnTyrAlaThGlyasnLeuProGlyCysSerPheSerIlePheLeu 178
 QY 184 TTGGCTTACTGTCCTGTCTGACCATTCAGACTCCGCTTGTAGGTGGCGACGTCGCC 243
 DB 179 LeuAlaLeuLeuSerCysleuthrIleProAlaSerAlaTyrGluValAlaArgValSer 198
 QY 244 GGGATGACCATGTCACGAAGAGCTGTCCAACTCAAGCATTTGTATGAGGACGCGAC 303
 DB 199 GlyIleTyrHisValThrAsnAspCysSerAsnSerSerIleValTyrGluAlaAlaAsp 218
 QY 304 ATGATCATGACACACCCCGGGTGGCTGCTGCTTGGGAGAACACTCTTCCCGCTGC 363
 DB 219 MetIleMetHisThrProGlyCysValProCysValArgGluGlyAsnSerSerArgCys 238
 QY 364 TTGGTACGCTGACCCCGACGCTGACAGCTAGAGAACGACGCTCCGACACTAGACATA 423
 DB 239 TrpValAlaLeuThrProThrLeuAlaAlaArgAsnAlaSerValProThrThrAlaIle 258
 QY 424 CGACGCCACGTCGATTTGCTGCTGGGGGCGCTTCTTGTCTCCGTATGACGTGGGG 483
 DB 259 ArgArgHisValAspLeuLeuValGlyAlaAlaAlaPheCysSerAlaMetTyrValGly 278
 QY 484 GATCTGGCGGATGCTCTCTCTCTCTCCAGCTGTTCACCATTCGGCTCCCGCGGCAT 543
 DB 279 AspLeuCysGlySerValPheLeuValSerClnLeuPheThrPheSerProAlaArgHis 298
 QY 544 GAGACGGTGCAGACTGCAATTTGCTCATCTATCCGGCCACATAGAGTACCGTATG 603
 DB 299 GluThrIleGlnAspCysAsnCysSerIleTyrProGlyHisValSerGlyHisArgMet 318
 QY 604 GCTTGGATATGATGATGAAGCTGG 627
 DB 319 AlaTrpAspMetMetMetAsnTrp 326

RESULT 9
 AAM98022 standard; protein; 3010 AA.
 ID AAM98022:
 AC AAM98022:
 XX 21-JUN-1999 (first entry) *done*
 DT 21-JUN-1999 (first entry) *not saved*
 XX Infectious hepatitis C virus genotype 1b strain HC-J4 protein.
 DE HCV; infectious clone; infection; diagnosis; therapy; vaccine;
 KM screening; assay; antiviral; virucide.
 XX Hepatitis C virus.
 OS Hepatitis C virus.
 XX W09904008-A2.
 PN 28-JAN-1999.
 XX 16-JUL-1998; 98MO-US14688.
 PF 27-JAN-1998; 98US-0014416.
 PR 18-JUL-1997; 97US-0053062.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Bukh J, Emerson SU, Purcell RH, Yanagi M;
 XX WPI: 1999-132252/11.
 DR N-PSDB; AAX24843.
 XX
 PT New isolated hepatitis C virus nucleic acids - used to develop
 PT products for the diagnosis, prevention and treatment of HCV
 PT infections and for developing screening assays
 PS Claim 2; Fig 14G-H; 126pp; English.

CC This protein is encoded by the infectious hepatitis C virus (HCV)
 CC genotype 1b strain HC-J4 genome (see AAX24833). HC-J4 was obtained
 CC from acute phase plasma of a chimpanzee infected with serum
 CC containing HC-J4/91. The infectious nucleic acid sequence can be
 CC used to produce chimeric genomes (see AAX24833) consisting of the
 CC open reading frames of infectious nucleic acid sequences of other
 CC genotypes (including genotypes 1-6) and subtypes (such as 1b, 2a,
 CC 2b, 2c, 3a, 4a-f, 5a and 6a) of HCV. The invention also relates to
 CC the introduction of mutations or deletions into infectious nucleic
 CC acid sequences in order to produce an attenuated HCV virus suitable
 CC for vaccine development. Infectious nucleic acid sequences can also
 CC be used to produce attenuated virus via passage in vitro or in vivo
 CC of the viruses produced by transfection of a host cell with the
 CC infectious nucleic acid sequence. Vaccines comprising one or more
 CC polypeptides made from the infectious nucleic acid sequence are
 CC used to immunize mammals, especially humans, against hepatitis C.
 CC The nucleic acid sequences can also be used to induce protective
 CC immunity against the virus. The nucleic acid sequences or their
 CC encoded proteases (e.g. NS3 protease) can additionally be used to
 CC develop screening assays to identify antiviral agents for HCV.

SQ Sequence 3010 AA:

Alignment Scores:
 Pred. No.: 5,46e-104 Length: 3010
 Score: 1081.00 Matches: 199
 Percent Similarity: 98.56% Conservative: 6
 Best Local Similarity: 95.67% Mismatches: 3
 Query Match: 93.59% Indels: 0
 DB: 20 Gaps: 0

US-09-899-303a-7 (1-633) x AAM98022 (1-3010)

QY 4 TTGGGTAAGGTCATGATACCTTACGTCGGGCTTCGCCGACCTCATGGGGTACATTCG 63
 DB 119 LeucllyysValIleapThrlleuthrCysGlyphealaapleumetGlyTyrIlePro 138
 QY 64 CTCGTCGGCCGCCCTAGGGGGTGTCCAGAGCCCTGGCGCATGGCGGTCTGTG 123
 DB 139 LeuValIleValProleuGlyGlyValIleAlaArgIleLeuAlaHISGlyValArgValVal 158
 QY 124 GAAGACGGCGGTGACATTCGACAGGGAATTTGGCTGTCTTCTTCTATCTTCCTC 183
 DB 159 GluaspIlyValasnTyrAlaThGlyasnLeuProGlyCysSerPheSerIlePheLeu 178
 QY 184 TTGGCTTACTGTCCTGTCTGACCATTCAGACTCCGCTTGTAGGTGGCGACGTCGCC 243
 DB 179 LeuAlaLeuLeuSerCysleuthrIleProAlaSerAlaTyrGluValAlaArgValSer 198
 QY 244 GGGATGACCATGTCACGAAGAGCTGTCCAACTCAAGCATTTGTATGAGGACGCGAC 303
 DB 199 GlyIleTyrHisValThrAsnAspCysSerAsnSerSerIleValTyrGluAlaAlaAsp 218
 QY 304 ATGATCATGACACACCCCGGGTGGCTGCTGCTTGGGAGAACACTCTTCCCGCTGC 363
 DB 219 MetIleMetHisThrProGlyCysValProCysValArgGluGlyAsnSerSerArgCys 238
 QY 364 TTGGTACGCTGACCCCGACGCTGACAGCTAGAGAACGACGCTCCGACACTAGACATA 423
 DB 239 TrpValAlaLeuThrProThrLeuAlaAlaArgAsnAlaSerValProThrThrAlaIle 258
 QY 424 CGACGCCACGTCGATTTGCTGCTGGGGGCGCTTCTTGTCTCCGTATGACGTGGGG 483

```

DB 259 ARGAGHNVASLSPLEUVALGLYTHRALAALAHPECYSERIALMETYVALGLY 278
QY 484 GATCTCTGGGATCTGTCCTCTCCAGCTGTCCAGCATATCGCCGCGGAT 543
DB 279 ASPLKCYSGLYSERILEPHEUVALSERGLNLEUPHETHPHESERPROARGHNS 298
QY 544 GAGACGGTCAGACATGCAATTGCTCAATCTATCCCGCCACATAACAGGTACCGTATG 603
DB 299 GLUTHVALGLNAPCYASNCYSERILETYPROGLYHISVALSERGLHISARGMET 318
QY 604 GCTTGGGATATGATGATGAACCTGG 627
DB 319 ALATRPSPMETMETMETASNTTP 326

RESULT 10
AAB31170
ID AAB31170 standard; Protein; 3010 AA.
XX
AC AAB31170;
XX
DE 02-APR-2001 (first entry)
XX
DE Amino acid sequence of a hepatitis C virus (HCV) clone genotype 1b.
XX
KM Chimeric virus; bovine viral diarrhoea virus; BVDV; hepatitis C virus;
KM HCV; vaccine; viral inhibitor; antiviral.
XX
OS Hepatitis C virus.
XX
PN WO200075352-A2.
XX
PD 14-DEC-2000.
XX
PE 02-JUN-2000; 2000WO-US15527.
XX
PR 04-JUN-1999; 99US-0137817.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Nam J, Bukh J, Emerson SU, Purcell RH;
XX
DR N-PSDB; AAC86939.
XX
PT WPI: 2001-071081/08.
XX
PT New nucleic acid comprising a chimeric bovine viral diarrhoea virus
PT genome in which the (non-)structural region has been replaced by
PT hepatitis C virus (HCV) genome useful for treating or preventing HCV
PT signs and symptoms -
XX
PS Disclosure: Fig 4G-H; 97p; English.
XX
XX The specification describes a nucleic acid comprising a chimeric virus
XX genome, specifically bovine viral diarrhoea virus (BVDV) genome in which
XX the (non-)structural region has been replaced by the (non-)structural
XX region of a hepatitis C virus (HCV) genome. The nucleic acids comprising
XX the chimeric virus and the chimeric virus are useful for identifying
XX cell lines capable of supporting the replication of these chimeric
XX viruses, in screening for neutralizing antibodies to HCV of different
XX genotypes, in the production of HCV-BVDV virions, for the development
XX of inactivated or attenuated vaccines to prevent HCV-BVDV in a mammal,
XX in studying the molecular properties of HCV indirectly in vitro, and in
XX identifying inhibitors of viral enzyme activity which would be useful
XX as antiviral agents. Formulations or compositions comprising the
XX chimeric virions may be used to treat or prevent the signs and symptoms
XX of HCV. The present sequence is encoded by a HCV clone, which is used
XX to construct chimeric nucleic acids of the invention.
XX
SQ Sequence 3010 AA;
XX
Alignment Scores: 5.46e-104 Length: 3010
Pred. No.: 1081.00 Matches: 199
Score:

```

```

Percent Similarity: 98.56%
Best Local Similarity: 95.67%
Query Match: 93.59%
DB: 22 Gaps: 0
us-09-899-303a-7 (1-633) x AAB31170 (1-3010)

QY 4 TTGGGTAAAGTCATGATGATCCCTTACGTCGGCTTCGCCAGCTCATGGGATATCCG 63
DB 119 LEUGLYLVVALILLEASPHRIEUTHRICYSGLYPHALASPLEUDETLYYRIEPRO 138
QY 64 CTCGTGGGCGCCCTTACGGGGGTGCTGCCAGACCCCTGGCGATGGCCGGGTCTG 123
DB 139 LEUVALGLYALAPROHEUDIGLYALALARGALALEALHISGLYVALARGYALLEU 158
QY 124 GAAGACGGCTGAATATGATCAACAGGATTTGCCGCTTCCTTCCTTCCTTCCTCC 183
DB 159 GLUASPGLYVALSNITRYALATHRGLYASNLEUPROGLYCYSSERPHESERILEPHEU 178
QY 184 TTGGCTTTACTGTCTCTGTCGACCATTCAGCTTCGCTTATGAGGTGGCAACGTCTCC 243
DB 179 LEUVALLEULEUSERCYSLEUTHRIEPROLASERIALTYGLVALARGSNVALSER 198
QY 244 GGGATGACCATGTCACGACGACACTGCTCCAACTAAGCATTTGTATGAGGACGGGAC 303
DB 199 GLYILETYRHISVALTHRSNAPCYSSERASNSEIRILEVALTYRGLUALALASAP 218
QY 304 ATGATCATGACACCCCGGGGTGGCGCTGCTGGGGAACAACTCTCCGCTCC 363
DB 219 VALLEMETHISTHPRDILYCYVALPROCYVALGINDGLYASNSEIRARGYCS 238
QY 364 TGGGTAGCCTCACCCCAAGCTCGACGTAGAAAGCCAGCGCTCCCACTAGCAATA 423
DB 239 TRPVALLALEUTHPROTHREUALLAARGAANALASERVALPROTHRTHRILE 258
QY 424 CGACGCCACGTGATTTGCTGTTGGGGCGCTCTTGTGTCCGATATGACGTGGG 483
DB 259 ARGARGHNSVALASPLEUVALGLYTHRALAALAHPECYSERIALMETYVALGLY 278
QY 484 GATCTCTGGGATCTGTCCTCTCCAGCTGTCCAGCATATCGCCGCGGAT 543
DB 279 ASPLKCYSGLYSERILEPHEUVALSERGLNLEUPHETHPHESERPROARGHNS 298
QY 544 GAGACGGTCAGACATGCAATTGCTCAATCTATCCCGCCACATAACAGGTACCGTATG 603
DB 299 GLUTHVALGLNAPCYASNCYSERILETYPROGLYHISVALSERGLHISARGMET 318
QY 604 GCTTGGGATATGATGATGAACCTGG 627
DB 319 ALATRPSPMETMETMETASNTTP 326

RESULT 11
AAE20477
ID AAE20477 standard; Protein; 3010 AA.
XX
AC AAE20477;
XX
DE 01-JUL-2002 (first entry)
XX
DE HCV-S1 full-length polyprotein.
XX
KM Nucleic acid construct; expression cassette; non-coding region; NCR;
KM untranslated region; UTR; anti-viral drug; drug resistance;
KM HCV-S1; Hepatitis C virus.
XX
OS Hepatitis C virus.
XX
PN WO200208447-A2.
XX
PD 31-JAN-2002.
XX
PF 20-JUL-2001; 2001WO-IL00669.
XX

```


CC develop screening assays to identify antiviral agents for HCV.

XX Sequence 3011 AA:

Alignment Scores:

Pred. No.:	5,47e-104	Length:	3011
Score:	1081.00	Matches:	199
Percent Similarity:	98.56%	Conservative:	6
Best Local Similarity:	95.67%	Mismatches:	3
Query Match:	93.59%	Indels:	0
DB:	20	Gaps:	0

US-09-899-303a-7 (1-633) x AAM98021 (1-3011)

```

QY 4 TTGGGTAGAGTCATGCATACCCCTTACGTGCGGCTTCCGACCTCATGGGATACATTCCG 63
DB 119 LeuGlyValAlaIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIlePro 138
QY 64 CTCGTGGGGGGCCCCCTAGGGGCTGCTGCACAGACCCCTGGCCATGGCGTCCGGTTCG 123
DB 139 LeuValGlyAlaProLeuGlyGlyAlaAlaArgAlaLeuAlaHisGlyValArgValLeu 158
QY 124 GAAGACGGGTGAAATATGACACAGGAAATTTGCTGCTCTTCTATCTCTC 183
DB 159 GluAspGlyValAsnTyrAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeu 178
QY 184 TTGGCTTAACTGTCGTGTGACCAATTCAGCTTCGCTTATGAGGTGGCAAGCTGTCC 243
DB 179 LeuAlaIleLeuSerCysLeuThrIleProAlaSerAlaTyrGlyValArgAsnValSer 198
QY 244 GGGATGACCATGTCAGAACGACGTGTCGAACCTCAAGCATTTGTATGAGCGAGGGGAC 303
DB 199 GlyIleTyrHisValThrAsnAspCysSerAsnSerSerIleValTyrGlyAlaAlaAsp 218
QY 304 ATGATCATGCACACCCCGGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363
DB 219 ValIleMetHisThrProGlyCysValProCysValGlnGlnGlyAsnSerSerArgCys 238
QY 364 TGGGTAGCGCTACACCCCGACGCTGCGAGCTAGAGACGCCAGCGCTCCCACTAGACATA 423
DB 239 TrpValAlaLeuThrProThrLeuAlaAlaArgAsnAlaSerValProThrThrIle 258
QY 424 CGACGCGACGTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 483
DB 259 ArgArgHisValAspLeuValGlyThrAlaAlaPheCysSerAlaMetGlyValGly 278
QY 484 GATCTGCGGATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 543
DB 279 AspLeuCysGlySerIlePheLeuValSerGlnLeuPheThrPheSerProArgArgHis 298
QY 544 GAGACGGGTGACGACTGCAATGCTATTCATCCGGCCACATAACAGGTGACCGGTATG 603
DB 299 GluThrValGlnAspCysAsnSerIleTyrProGlyHisValSerGlyHisAlaMet 318
QY 604 GCTTGGGATATGATGAAGCTGG 627
DB 319 AlaTrpAspMetMetIleAsnTrp 326

```

RESULT 13
AAB59174
ID AAB59174 standard; Protein; 3010 AA.

AC AAB59174;

DT 21-MAR-2001 (first entry)

DE Protein encoded by infectious Hepatitis C virus 1b genotype.

KM GBV-B; hepatitis C virus; HCV; vaccine.

OS Hepatitis C virus.

XX WO200075337-A1.

```

XX 14-DEC-2000.
PD 02-JUN-2000; 2000WO-US15293.
XX 04-JUN-1999; 99US-0137694.
PR (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA Bukh J, Yanagi M, Emerson SU, Purcell RH;
PI WPI; 2001-091214/10.
XX
DB New infectious nucleic acids of the GB virus-B clone, useful for
PT indirectly studying the molecular properties of hepatitis C virus (HCV)
PT and in developing vaccines and therapeutics for HCV
XX
PS Disclosure; Fig 7; 96pp; English.
XX
CC The present invention relates to GB virus-B. The nucleic acid molecules
CC of the invention are useful for indirectly studying the molecular
CC properties of hepatitis C virus (HCV). The infectious nucleic acid
CC sequence of the GB virus-B clone and the HCV/GBV-B chimeras may be used
CC in the development of vaccines and therapeutics for HCV.
XX
SQ Sequence 3010 AA:

```

Alignment Scores:

Pred. No.:	2.32e-103	Length:	3010
Score:	1075.00	Matches:	198
Percent Similarity:	98.08%	Conservative:	6
Best Local Similarity:	95.19%	Mismatches:	4
Query Match:	93.07%	Indels:	0
DB:	22	Gaps:	0

US-09-899-303a-7 (1-633) x AAB59174 (1-3010)

```

QY 4 TTGGGTAGAGTCATGCATACCCCTTACGTGCGGCTTCCGACCTCATGGGATACATTCCG 63
DB 119 LeuGlyValAlaIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIlePro 138
QY 64 CTCGTGGGGGGCCCCCTAGGGGCTGCTGCACAGACCCCTGGCCATGGCGTCCGGTTCG 123
DB 139 LeuValGlyAlaProLeuGlyGlyAlaAlaArgAlaLeuAlaHisGlyValArgValLeu 158
QY 124 GAAGACGGGTGAAATATGACACAGGAAATTTGCTGCTCTTCTATCTCTC 183
DB 159 GluAspGlyValAsnTyrAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeu 178
QY 184 TTGGCTTAACTGTCGTGTGACCAATTCAGCTTCGCTTATGAGGTGGCAAGCTGTCC 243
DB 179 LeuAlaIleLeuSerCysLeuThrIleProAlaSerAlaTyrGlyValArgAsnValSer 198
QY 244 GGGATGACCATGTCAGAACGACGTGTCGAACCTCAAGCATTTGTATGAGCGAGGGGAC 303
DB 199 GlyIleTyrHisValThrAsnAspCysSerAsnSerSerIleValTyrGlyAlaAlaAsp 218
QY 304 ATGATCATGCACACCCCGGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363
DB 219 ValIleMetHisThrProGlyCysValProCysValGlnGlnGlyAsnSerSerArgCys 238
QY 364 TGGGTAGCGCTACACCCCGACGCTGCGAGCTAGAGACGCCAGCGCTCCCACTAGACATA 423
DB 239 TrpValAlaLeuThrProThrLeuAlaAlaArgAsnAlaSerValProThrThrIle 258
QY 424 CGACGCGACGTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 483
DB 259 ArgArgHisValAspLeuValGlyThrAlaAlaPheCysSerAlaMetGlyValGly 278
QY 484 GATCTGCGGATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 543
DB 279 AspLeuCysGlySerIlePheLeuValSerGlnLeuPheThrPheSerProArgArgHis 298

```

```

OY 544 GAGACGGTCGACGACTGCAATTGCTCAATCTATCCGCCACATACAGGTCACCGTATG 603
DB 299 GltthrValGlnAspCysAsnCySerIleTyrProGlyHisValSerGlyHisArgMet 318
OY 604 GCTTGGATATGATGATGAACCTGG 627
DB 319 AlatrPaspMetMetMetAsnTirP 326

RESULT 14
AAR25854
ID AAR25854 standard; Protein; 416 AA.
XX
AC AAR25854;
XX
DT 21-JAN-1993 (first entry)
XX
DE HCV polypeptide 1.
XX
KM Recombinant vector; E. coli; diagnostic; reagent; type C hepatitis.
XX
OS Hepatitis C virus.
XX
PN JP04179482-A.
XX
PD 26-JUN-1992.
XX
PF 11-NOV-1990; 90JP-0304417.
XX
PR 11-NOV-1990; 90JP-0304417.
XX
PA (TOKU ) TOKUYAMA SODA KK.
XX
DR WPI; 1992-263663/32.
XX
N-PSDB; AAQ26981.
XX
PT Hepatitis C virus antigen expressed as recombinant in E.coli -
PT useful for diagnosis of hepatitis C virus infection
XX
PS Claim 1; Page 1-2; 66pp; Japanese.
XX
CC The sequences given in AAR25854-74 are hepatitis C virus proteins.
CC The genes encoding these proteins can each be used to prepare
CC recombinant vectors by ligating the gene of interest in to a vector
CC to be expressed in E. coli. These polypeptides are useful as
CC diagnostic reagents for type C hepatitis and they may be produced
CC efficiently by recombinant methods.
XX
SO Sequence 416 AA:

Alignment Scores:
Pred. No.: 1.88e-103 Length: 416
Score: 1073.00 Matches: 198
Percent Similarity: 98.56% Conservative: 7
Best Local Similarity: 95.19% Mismatches: 3
Query Match: 92.90% Indels: 0
DB: 13 Gaps: 0

US-09-899-303a-7 (1-633) x AAR25854 (1-416)
OY 4 TTGGGTAAAGTCAATGATACCTTACGTCGGGCTTCCGCCACCTCATGGGTCATTCG 63
DB 24 LeuGlyValIleAspTrpLeuThrCysGlyPheAlaAspLeuMetGlyTyrIlePro 43
OY 64 CTGCTGCGCCGCCCTAGGGGTCGTCACAGAGCCCTGGCGCATGGCGTCCGGGTTCTG 123
DB 44 LeuValGlyAlaProLeuGlyGlyAlaAlaArgAlaLeuAlaHisGlyValAlaGValLeu 63
OY 124 GAAGAGCGCGTGAACATGACAGAGGAATTCGCTGGTTCCTTCTATCTTCCTC 183
DB 64 GluAspIleValAsnTyrAlaThrGlyAsnLeuProGlyAsnSerPheSerIlePheLeu 83
OY 184 TTGGCTTACTGTCCTGTCGACCATTCGAGCTTCGCTTATGAGTGGCGCAACGTCTC 243

```

```

DB 84 LeuAlaLeuLeuSerCysLeuThrIleProAlaSerAlaTyrGluValArgAsnValSer 103
OY 244 GCGATGTACCAATGTACAGAAAGACTGCTCCAACTCAAGCAATGTGTATGAGCGAC 303
DB 104 GlyIleTyrHisValThrAsnAspCysSerAsnSerIleValTyrGluAlaAlaAsp 123
OY 304 ATGATCATGCACACCCCCGGGTGCGTCCCTGCGTGGAGAAACAATCTTCCCGTGC 363
DB 124 MetIleMetHisThrProGlyCysValProCysValArgGluSerAsnPheSerArgCys 143
OY 364 TTGGTACGCTCAACCCCGACGCTCGCAGTCAGAAACCCGACGTCCTACGACATA 423
DB 144 TrpAlaAlaLeuThrProThrLeuAlaAlaArgAsnSerIleProThrThrIle 163
OY 424 CGACGCCACGCTGATTTGCTGCTGGGGCGGCGCTTCTTCCCGTATGACGTGGG 483
DB 164 ArgThrHisValAspLeuLeuValGlyAlaAlaAlaLeuCysSerAlaMetTyrValGly 183
OY 484 GATCTCTGCGGATCTGCTTCCTGCTCCAGCTGTTCACATCTCCGCTCGCGGCAT 543
DB 184 AspLeuCysGlySerValPheLeuValSerGlnLeuPheThrPheSerProArgArgTyr 203
OY 544 GAGACGGTCGACGACTGCAATTGCTCAATCTATCCGCCACATACAGGTCACCGTATG 603
DB 204 GltthrValGlnAspCysAsnCySerIleTyrProGlyHisValSerGlyHisArgMet 223
OY 604 GCTTGGATATGATGATGAACCTGG 627
DB 224 AlatrPaspMetMetMetAsnTirP 231

RESULT 15
AAR25876
ID AAR25876 standard; Protein; 445 AA.
XX
AC AAR25876;
XX
DT 21-JAN-1993 (first entry)
XX
DE HK1.
XX
KM Recombinant vector; E. coli; diagnostic; reagent; type C hepatitis.
XX
OS Hepatitis C virus.
XX
FH Key Location/Qualifiers
FT protein 25..441
FT /*tag=
FT /*note= "Sequence AAR25854"

JP04179482-A.
PN
PD 26-JUN-1992.
XX
PF 11-NOV-1990; 90JP-0304417.
XX
PR 11-NOV-1990; 90JP-0304417.
XX
PA (TOKU ) TOKUYAMA SODA KK.
XX
DR WPI; 1992-263663/32.
XX
N-PSDB; AAQ27003.
XX
PT Hepatitis C virus antigen expressed as recombinant in E.coli -
PT useful for diagnosis of hepatitis C virus infection
XX
PS Disclosure; Fig 2; 66pp; Japanese.
XX
CC The sequences given in AAR25876-95 are encoded by the claimed hepatitis
CC C virus genes of the invention which have been inserted into an E.
CC coli vector. These polypeptides are useful as diagnostic reagents
CC for type C hepatitis and they may be produced efficiently by
CC recombinant DNA techniques.
XX

```

SQ Sequence 445 AA:

Alignment Scores:

Pred. No.:	1.93e-103	Length:	445
Score:	1073.00	Matches:	198
Percent Similarity:	98.56%	Conservative:	7
Best Local Similarity:	95.19%	Mismatches:	3
Query Match:	92.90%	Indels:	0
DB:	13	Gaps:	0

US-09-899-303A-7 (1-633) x AAR25876 (1-445)

```
OY      4  TTGGGTAAGTCATGATCCCTTACGTGCGGCTTGCCGACCTCATGGGGTACATTCGG 63
      |||
DB     49  LeuGlyLysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIlePro 68
OY     64  CTCGTGCGGCGCCCTTACGGGGTCTGCGCAGAGCCCTGGCCATGAGCGTCCGGGTTCTG 123
      |||
DB     69  LeuValGlyAlaProLeuGlyGlyAlaAlaArgAlaLeuAlaHisGlyValArgValLeu 88
OY     124 GAAGACGGCGTGAACATGATCAACAGGGAATTTGCGTGTCTTCTCTATCTTCTC 183
      |||
DB     89  GluAspGlyValAsnTyrAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeu 108
OY     184 TTGGGCTTACTGTCCTGTCTGACACATTCAGCTTCGGCTTATGAGTGGCAACGTGTC 243
      |||
DB     109 LeuAlaLeuLeuSerCysLeuThrIleProAlaSerAlaTyrGlyValArgAsnValSer 128
OY     244 GGGATGTACCATGTATCGAAGACGCTGCTCAACTCAAGCATTTGTATGAGCGACGGAC 303
      |||
DB     129 GlyIleTyrHisValThrAsnAspCysSerAsnSerIleValTyrGluAlaAlaAsp 148
OY     304 ATGATCATGACACACCCCGGGTGCCTGCGCTTCGTTGCGGAGAACACTTTCGCCGTGC 363
      |||
DB     149 MetIleMetHisThrProGlyCysValProCysValArgIleSerAsnPheSerArgCys 168
OY     364 TGGGTAGCCCTCACCCCGCTCGAGCTAGAGACGCCAGCGTCCCGCACTACGACATA 423
      |||
DB     169 TrpValAlaLeuThrProThrLeuAlaAlaArgAsnSerIleProThrThrIle 188
OY     424 CGACGCCACGTGATTTGCTGCTGGGGGGGCTTCTGTTCCGCTATGTACGTGGGG 483
      |||
DB     189 ArgAlaGlyHisValAspLeuLeuValGlyAlaAlaAlaLeuCysSerAlaMetTyrValGly 208
OY     484 GATCTCTGCGGATCTGCTTCTCTGCTCCAGCTGTTACCACTTCGGCTCGCCGGCAT 543
      |||
DB     209 AspLeuCysGlySerValPheLeuValSerGlnLeuPheThrPheSerProArgArgTyr 228
OY     544 GAGACGGTGCAGAGCTGCAATGCTAATCTATCCGGCCACATAACAGAGTCAACGCTATG 603
      |||
DB     229 GluThrValGlnAspCysAsnCysSerIleTyrProGlyHisValSerGlyHisArgMet 248
OY     604 GCTTGGGATATGATGATGAAGTGG 627
      |||
DB     249 AlaTrpAspMetMetIasnTrp 256
```

Search completed: February 12, 2003, 08:58:46
Job time : 51 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: February 12, 2003, 08:54:59 ; Search time 14 Seconds

(without alignments)
2660.673 Million cell updates/sec

Title: US-09-899-303A-7

Perfect score: 1155
Sequence: 1 ATGTGCGTAAGCTCATCGA.....TGATGATGAAGCTGTAATAG 633

Scoring table:

BL0SUM62	
Xgapop 10.0 , Xgapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 262574 segs, 29422922 residues

Total number of hits satisfying chosen parameters: 525148

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xip
-O/cgn2_1/uspto.spool/US09899303/rnat.12022003.085450.28314/app.query.fasta.1.775
-DB=Issued Patents_AA -QPM=fastan -SUFFIX=n2p.ra1 -MTNMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bl0sum62 -TRANS=human40.cdl
-LIST=45 -DOCLIN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTPM=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09899303 -CGN1_1.12 -Guanat.12022003.085450.28314 -NCPU=6 -ICPU=3
-NO.XLPXY -NO.MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-MARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

Issued Patents_AA: *
1: /cgn2_6/prodata/1/iaa/5A.COMB.pep: *
2: /cgn2_6/prodata/1/iaa/5B.COMB.pep: *
3: /cgn2_6/prodata/1/iaa/6A.COMB.pep: *
4: /cgn2_6/prodata/1/iaa/6B.COMB.pep: *
5: /cgn2_6/prodata/1/iaa/PCTUS.COMB.pep: *
6: /cgn2_6/prodata/1/iaa/backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1118	96.8	209	4	US-08-612-973-8
2	1118	96.8	209	4	US-08-927-597-8
3	1113	96.4	263	4	US-08-612-973-6
4	1113	96.4	263	4	US-08-927-597-6
5	1109	96.0	210	4	US-08-612-973-14
6	1109	96.0	210	4	US-08-927-597-14
7	1108	95.9	632	4	US-08-612-973-48
8	1108	95.9	632	4	US-08-927-597-48
9	1108	95.9	809	4	US-08-612-973-50
10	1108	95.9	809	4	US-08-927-597-50
11	1084	93.9	340	1	US-08-462-195-2
12	1084	93.9	340	2	US-08-636-883-2

13	1084	93.9	340	3	US-09-127-829-2	Sequence 2, Appl1
14	1081	93.6	3011	4	US-09-014-416-3	Sequence 3, Appl1
15	1081	93.6	3011	4	US-09-014-416-5	Sequence 5, Appl1
16	1065	92.2	340	1	US-08-462-195-4	Sequence 4, Appl1
17	1065	92.2	340	2	US-08-636-883-4	Sequence 4, Appl1
18	1065	92.2	340	3	US-09-127-829-4	Sequence 4, Appl1
19	1060	91.8	389	1	US-08-324-977-4	Sequence 4, Appl1
20	1060	91.8	389	1	US-08-324-977-10	Sequence 10, Appl1
21	1060	91.8	389	2	US-08-384-616-4	Sequence 4, Appl1
22	1060	91.8	389	2	US-08-384-616-10	Sequence 10, Appl1
23	1060	91.8	389	2	US-08-904-686A-4	Sequence 4, Appl1
24	1060	91.8	389	2	US-08-904-686A-10	Sequence 10, Appl1
25	1060	91.8	389	4	US-09-315-850-4	Sequence 4, Appl1
26	1060	91.8	389	4	US-09-315-850-12	Sequence 12, Appl1
27	1060	91.8	2013	1	US-08-324-977-12	Sequence 12, Appl1
28	1060	91.8	2013	2	US-08-384-616-12	Sequence 12, Appl1
29	1060	91.8	2013	2	US-08-904-686A-12	Sequence 12, Appl1
30	1060	91.8	2013	4	US-09-315-850-12	Sequence 12, Appl1
31	1060	91.8	3010	1	US-08-324-977-2	Sequence 2, Appl1
32	1060	91.8	3010	1	US-08-324-977-14	Sequence 14, Appl1
33	1060	91.8	3010	2	US-08-384-616-2	Sequence 2, Appl1
34	1060	91.8	3010	2	US-08-384-616-14	Sequence 14, Appl1
35	1060	91.8	3010	2	US-08-904-686A-2	Sequence 2, Appl1
36	1060	91.8	3010	2	US-08-904-686A-14	Sequence 14, Appl1
37	1060	91.8	3010	4	US-09-315-850-2	Sequence 2, Appl1
38	1060	91.8	3010	4	US-09-315-850-14	Sequence 14, Appl1
39	1006	87.1	967	1	US-08-188-281B-13	Sequence 13, Appl1
40	1006	87.1	967	5	PCT-US94-07280-13	Sequence 13, Appl1
41	1006	87.1	967	5	PCT-US95-01087-13	Sequence 13, Appl1
42	1006	87.1	1648	1	US-08-188-281B-12	Sequence 12, Appl1
43	1006	87.1	1648	5	PCT-US94-07280-12	Sequence 12, Appl1
44	1006	87.1	1648	5	PCT-US95-01087-12	Sequence 12, Appl1
45	1006	87.1	3011	1	US-08-188-281B-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-08-612-973-8
; Sequence 8, Application US/08612973
; Patent No. 6150134
GENERAL INFORMATION:
APPLICANT: MAERTENS, GEERT
APPLICANT: BOSMAN, FONS
APPLICANT: DE MARTYNOFF, GUY
APPLICANT: BUYSSE, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESSES:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612,973
FILING DATE: 11-MAR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 8:

```

;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 209 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-612-973-8

Alignment Scores:
Pred. No.: 8.35e-111 Length: 209
Score: 1118.00 Matches: 209
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 96.80% Indels: 0
DB: 4 Gaps: 0

US-09-899-303A-7 (1-633) x US-08-612-973-8 (1-209)
QY 1 ATGTTGGTAAAGTCAATGATACCTTACGTGGCGCTTGCCGACCTCATGGGGTACATT 60
DB 1 MetLeuGlYlYsValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIle 20
QY 61 CGGCTGTGGGGGCCCCCTAGGGGGTGCCTCCAGAGCCCTGGCGCATGGCGTCCGGATT 120
DB 21 ProlLeuValIGlYAlaProlLeuGlyGlyAlaAlaArgAlaLeuAlaHisGlyValArgVal 40
QY 121 CTGGGAAGAGGGGTGAACATGCAACAGGGAATTTCCCTGCTTCTTCTATCTTC 180
DB 41 LeuGluAspGlyValAsnGlyAlaThrGlyAsnLeuProlGlyCysSerPheSerIlePhe 60
QY 181 CTTTGGCTTTACTGTCCTGTGACCATTCACGCTTCCGCTATGAGGTGGCCACAGT 240
DB 61 LeuLeuAlaLeuLeuSerGlyLeuThrIleProlAlaSerAlaTyrGluValArgAsnVal 80
QY 241 TCCGGGATGTACCATGTGTCAAGACGACTGTCTCCAACTCAAGCATTTGTATGAGCGAG 300
DB 81 SerGlyMetTyrHisValIleHisAsnAspCysSerAsnSerSerIleValIyrGluAlaAla 100
QY 301 GACATGATATGACACACCCCGGGTGGCTGGCTGCGCTTCCGGAGAACACTTTCCCG 360
DB 101 AspMetIleMetHisThrProlGlyCysValProCysValArgGluAsnAsnSerSerArg 120
QY 361 TCGTGGGTAGCGCTCAACCCCGACGCTCGAGCTAGGAGCCGACGCTCCCACTAGAC 420
DB 121 CysTrpValAlaLeuThrProThrIleuAlaAlaArgAsnAlaSerValProThrThr 140
QY 421 ATACAGACGCGACGTGATTTGCTGCTGGGGGCGCTTCTTGTTCCTATGACGTG 480
DB 141 IleArgGlnHisValAspLeuLeuValIGlYAlaAlaAlaPheCysSerAlaMetGlyVal 160
QY 481 GGGGATCTGTGGGATCTGTCTTCCCTGCTGCCAGCTGTTCACCATCTGCGCTCGCG 540
DB 161 GlysAspLeuCysGlySerValPheLeuValSerGlnLeuPheThrIleSerProArgArg 180
QY 541 CATGAGACGCTGACGATGCAATGCTCAATCTATCCGGGCGACATACAGGTACCGGT 600
DB 181 HisGluThrValIGlYAlaAspCysAsnCysSerIleTyrProlGlnHisIleThrGlnHisArg 200
QY 601 ATGCGTTGGGATATGATGATGAACGTG 627
DB 201 MetAlaIleTrpAspMetMetMetAsnTrp 209

RESULT 2
US-08-927-597-8
; Sequence 8, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: NAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
```

```

;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P. C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597
; FILING DATE:
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 209 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-927-597-8

Alignment Scores:
Pred. No.: 8.35e-111 Length: 209
Score: 1118.00 Matches: 209
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 96.80% Indels: 0
DB: 4 Gaps: 0

US-09-899-303A-7 (1-633) x US-08-927-597-8 (1-209)
QY 1 ATGTTGGTAAAGTCAATGATACCTTACGTGGCGCTTGCCGACCTCATGGGGTACATT 60
DB 1 MetLeuGlYlYsValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIle 20
QY 61 CGGCTGTGGGGGCCCCCTAGGGGGTGCCTCCAGAGCCCTGGCGCATGGCGTCCGGATT 120
DB 21 ProlLeuValIGlYAlaProlLeuGlyGlyAlaAlaArgAlaLeuAlaHisGlyValArgVal 40
QY 121 CTGGGAAGAGGGGTGAACATGCAACAGGGAATTTCCCTGCTTCTTCTATCTTC 180
DB 41 LeuGluAspGlyValAsnGlyAlaThrGlyAsnLeuProlGlyCysSerPheSerIlePhe 60
QY 181 CTTTGGCTTTACTGTCCTGTGACCATTCACGCTTCCGCTATGAGGTGGCCACAGT 240
DB 61 LeuLeuAlaLeuLeuSerGlyLeuThrIleProlAlaSerAlaTyrGluValArgAsnVal 80
QY 241 TCCGGGATGTACCATGTGTCAAGACGACTGTCTCCAACTCAAGCATTTGTATGAGCGAG 300
DB 81 SerGlyMetTyrHisValIleHisAsnAspCysSerAsnSerSerIleValIyrGluAlaAla 100
QY 301 GACATGATATGACACACCCCGGGTGGCTGGCTGCGCTTCCGGAGAACACTTTCCCG 360
DB 101 AspMetIleMetHisThrProlGlyCysValProCysValArgGluAsnAsnSerSerArg 120
QY 361 TCGTGGGTAGCGCTCAACCCCGACGCTCGAGCTAGGAGCCGACGCTCCCACTAGAC 420
DB 121 CysTrpValAlaLeuThrProThrIleuAlaAlaArgAsnAlaSerValProThrThr 140
QY 421 ATACAGACGCGACGTGATTTGCTGCTGGGGGCGCTTCTTGTTCCTATGACGTG 480
```

Db 141 ILeArGrhISValAsPLeuLeuValGlyAlaAlaIaPheCysSerAlaMetLyrVal 160
QY 481 GGGATGTCGCGGATGTCCTTCGTCGCCAGCTGTCACCATCTCCGCGCG 540
Db 161 GlyAsPLeuCysGlySerValPheLeuValSerGlnLeuPheThrIleSerProArgArg 180
QY 541 CATGAGCGGTGACGAGACTGCATTCATTCATCCCGGCCACATMAAGATCACCCT 600
Db 181 HIsGlnThrValGlnAspCysAsnCysSerIleLyrProGlyHisIleThrGlyHisArg 200
QY 601 ATGCGCTGGGATATGATGATGACTGG 627
Db 201 MetAlaTrpAspMetMetMetLsnTrp 209
RESULT 3
US-08-612-973-6
; Sequence 6, Application US/08612973
; Patent No. 6130134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHAYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 263 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-612-973-6
Alignment Scores:
Pred. No.: 3.13e-110 Length: 263
Score: 1113.00 Matches: 207
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 99.04% Mismatches: 0
Query Match: 96.36% Indels: 0
DB: 4 Gaps: 0
US-09-899-303a-7 (1-633) x US-08-612-973-6 (1-263)
QY 1 ATGTTGGGTAAAGTCAATGATACCTTACGTCGCGCTTCGCCGACCTCATGGGGTACATT 60
Db 1 MetLeuGlyGlyValIleAspThrLeuThrCysGlyPheAlaAspLeuValGlyTyrIle 20
QY 61 CCGCTCGTCGGCGCCCTTAAGGGGTGCTGCCAGAGCCCTGGCGCATGGCGTCCGGGTT 120

Db 21 ProLeuValGlyAlaProLeuGlyGlyAlaAlaIaPheAlaIaHisGlyValArgVal 40
QY 121 CTGGAAGCGCGGAGACTATGCAACAGGGGAATTGGCTGGCTCTTCTCATCTTC 180
Db 41 LeuGlnAspGlyValAsnTrpAlaThrIleGlyAsnLeuProGlyCysSerPheSerIlePhe 60
QY 181 CTCCTGGCTTACTGTCTCTGTCTGACCATTCACACTTCGCTTATGAGAGTCCGACGTG 240
Db 61 LeuLeuAlaLeuLeuSerCysLeuThrValProAlaSerAlaTyrGluValArgAsnVal 80
QY 241 TCCGGAGATGACATGTCAGACAGACTGTCTCCACATTCACATTCATGATGATAGCGACGG 300
Db 81 SerGlyMetLyrHisValThrAsnAspCysSerAsnSerIleValTyrGluAlaAla 100
QY 301 GACATGATCATGACACACCCCGCGGTGGCTGCCCTTCGGGAGAACACTTCTCCCGC 360
Db 101 AspMetIleMetHisThrProGlyCysValProCysValArgGlnAsnAsnSerSerArg 120
QY 361 TCGTGGGTAGCGCTACACCCCGCACGCTCGACGTAAGAACGCGACGCTCCACATAGACA 420
Db 121 CysTrpAlaIaLeuThrProThrLeuAlaIaIaArgAsnAlaSerValProThrThr 140
QY 421 ATACGAGCCGACGTCGATTCCTGTTGGGGCGCGCTGCTTCTGTCCGTATGTACGTG 480
Db 141 ILeArGrhISValAsPLeuLeuValGlyAlaAlaIaPheCysSerAlaMetLyrVal 160
QY 481 GGGATGTCGCGGATGTCCTTCGTCGCCAGCTGTCACCATCTCCGCGCGCG 540
Db 161 GlyAsPLeuCysGlySerValPheLeuValSerGlnLeuPheThrIleSerProArgArg 180
QY 541 CATGAGCGGTGACGAGACTGCATTCATTCATCCCGGCCACATMAAGATCACCCT 600
Db 181 HIsGlnThrValGlnAspCysAsnCysSerIleLyrProGlyHisIleThrGlyHisArg 200
QY 601 ATGCGCTGGGATATGATGATGACTGG 627
Db 201 MetAlaTrpAspMetMetMetLsnTrp 209
RESULT 4
US-08-927-597-6
; Sequence 6, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHAYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205

```

REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 263 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-927-597-6

Alignment Scores:
Pred. No.: 3,13e-110
Score: 1113.00
Percent Similarity: 100.00%
Best Local Similarity: 99.04%
Query Match: 96.36%
Matches: 263
Conservative: 2
Mismatch: 0
Indels: 0
Gaps: 0

US-09-899-303A-7 (1-633) x US-08-927-597-6 (1-263)
QY 1 ANCTGGGTAAAGTCATCGATACCTTACGTGGGCTGCCGACCTCATGGGTCATTT
Db 1 MetLeuGlyLysValIleAspThrLeuThrCysGlyPheAlaAspLeuValGlyTyrIle 20
QY 61 CCGCTGTCGGCGCCGCCCTAGGGGGTCTGCCAGAGCCCTGGCGATGGCGCTCCGGGTT
Db 21 ProlLeuValIGlyAlaProlLeuGlyGlyAlaAlaArgAlaLeuAlaHisGlyValArgVal 40
QY 121 CTGGAAGAGCGGTGAACATATGCAACAGGAATTCCTGCTGCTTCTTCTATCTTC
Db 41 LeuGluAspGlyValAsnTyrAlaThrGlyAsnLeuProlGlyCysSerPheSerIlePhe 60
QY 181 CTCTGGCTTTACTGTCCTGTGTCAGCATTCCTCCATTCAGTGGGCGCAACGTG
Db 61 LeuLeuAlaLeuLeuSerGlyLeuThrValProlAlaSerAlaTyrGluValArgAsnVal 80
QY 241 TCCGGGATGTACCATGTCAAGAACAGACTGCTCCAACTCAAGCATTTGTATGAGGAGCG
Db 81 SerGlyMetTyrHisValThrAsnAspCysSerAsnSerSerIleValTyrGluAlaAla 100
QY 301 GACATGATCATGCACACCCCGGGGTGCTGGCTGGGTTCCGGAGAACACTTTCCTCCG
Db 101 AspMetIleMetHisThrProGlyCysValProCysValArgGluAsnAsnSerSerArg 120
QY 361 TCGTGGTACGCTCACCCCGGAGCTCGAGCTAGAGAACGCCAGCTCCCACTAGACAGA
Db 121 CysTrpValAlaLeuThrProThrIleuAlaAlaArgAsnAlaSerValProThrThrThr 140
QY 421 ATACGACCCGACGTCGATTTGCTGCTGGGGCGGCTGCTTCTGTTCGCTATGTACGTG
Db 141 IleArgArgHisValAspLeuLeuValIGlyAlaAlaAlaPheCysSerAlaMetTyrVal 160
QY 481 GGGGATCTGTGGGATCTGCTTCTCCGCTCCAGCTTTCACCATCTCCGCTCCGCGG
Db 161 GlyAspLeuCysGlySerValPheLeuValSerGlnLeuPheThrIleSerProAlaArg 180
QY 541 CATTGACCGCTGACGAGCTGCAATTGCTCAATCTATCCGGCCATTAACAGGTACACCGT
Db 181 HisGluThrValGlnAspCysAsnSerIleTyrProGlyHisIleThrGlyHisArg 200
QY 601 ATGGCTGGGATGTATGATGAACCTGG 627
Db 201 MetAlaTrpAspMetMetAsnTrp 209

```

```

APPLICANT: DE MARTYNOFF, GUY
APPLICANT: BOYSE, MARIE-ANGE
TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESS: NIXON & VANDERHUYE P. C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612,973
FILING DATE: 11-MAR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 210 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-612-973-14

Alignment Scores:
Pred. No.: 7.59e-110
Score: 1109.00
Percent Similarity: 99.52%
Best Local Similarity: 98.56%
Query Match: 96.02%
Matches: 210
Conservative: 2
Mismatch: 1
Indels: 0
Gaps: 0

US-09-899-303A-7 (1-633) x US-08-612-973-14 (1-210)
QY 1 ANCTGGGTAAAGTCATCGATACCTTACGTGGGCTGCCGACCTCATGGGTCATTT
Db 1 MetLeuGlyLysValIleAspThrLeuThrCysGlyPheAlaAspLeuValGlyTyrIle 20
QY 61 CCGCTGTCGGCGCCGCCCTAGGGGGTCTGCCAGAGCCCTGGCGATGGCGCTCCGGGTT
Db 21 ProlLeuValIGlyAlaProlLeuGlyGlyAlaAlaArgAlaLeuAlaHisGlyValArgVal 40
QY 121 CTGGAAGAGCGGTGAACATATGCAACAGGAATTCCTGCTGCTTCTTCTATCTTC
Db 41 LeuGluAspGlyValAsnTyrAlaThrGlyAsnLeuProlGlyCysSerPheSerIlePhe 60
QY 181 CTCTGGCTTTACTGTCCTGTGTCAGCATTCCTCCATTCAGTGGGCGCAACGTG
Db 61 LeuLeuAlaLeuLeuSerGlyLeuThrValProlAlaSerAlaTyrGluValArgAsnVal 80
QY 241 TCCGGGATGTACCATGTCAAGAACAGACTGCTCCAACTCAAGCATTTGTATGAGGAGCG
Db 81 SerGlyMetTyrHisValThrAsnAspCysSerAsnSerSerIleValTyrGluAlaAla 100
QY 301 GACATGATCATGCACACCCCGGGGTGCTGCCAGAGCTCCCACTCAAGCATTTGTCCCGC
Db 101 AspMetIleMetHisThrProGlyCysValProCysValArgGluAsnAsnSerSerArg 120
QY 361 TCGTGGTACGCTCACCCCGGAGCTCGAGCTAGAGAACGCCAGCTCCCACTAGACAGA
Db 121 CysTrpValAlaLeuThrProThrIleuAlaAlaArgAsnAlaSerIleProThrThrThr 140

```


[illegible]

NAME: BYRNE, THOMAS E.
 REGISTRATION NUMBER: 32,205
 REFERENCE/DOCKET NUMBER: 1487-10
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 816-4000
 TELEFAX: (703) 816-4100
 INFORMATION FOR SEQ ID NO: 48:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 692 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-612-973-48

 Alignment Scores:
 Pred. No.: 1.59e-109 Length: 692
 Score: 1108.00 Matches: 206
 Percent Similarity: 100.00% Conservative: 2
 Best Local Similarity: 99.04% Mismatches: 0
 Query Match: 95.93% Indels: 0
 DB: 4 Gaps: 0

 US-09-899-303A-7 (1-633) x US-08-612-973-48 (1-692)
 QY 4 TTGGGTAAGTCAATCGATACCTTACGTGGCGCTTCCGCGACCTCATGGGCTACATTCCG 63
 Db 2 LeuG1yLysVal11IleAspThrLeuThrCysG1yPheAlaAspLeuValG1yTyr1IlePro 21
 QY 64 CTCGTGGGCGCCCGCCAGGCGGGGCGTCCAGAGCCCTGGCGCATGGCGCGTTCGTTCTG 123
 Db 22 LeuValG1yAlaProLeuG1yG1yAlaAlaArgAlaLeuAlaHisG1yValAlaGValLeu 41
 QY 124 GAAGACGGCGGTGAATGACAGACAGGGAATTTGCGTGGTTCCTCTTCTATCTTCCTC 183
 Db 42 GluAspG1yAlaAsnTyrAlaThrG1yAsnLeuProG1yLysSerPheSer1IlePheLeu 61
 QY 184 TTGGCTTTACTGCTCCGTGACCAATCCAGCTTCGCTTATGAGTGCGCCCAACGCTGCC 243
 Db 62 LeuAlaLeuLeuSerCysLeuThrValProAlaSerAlaTyrG1yValAlaGAsnValSer 81
 QY 244 GGGATGTACCATGTCAAGACGACTGCTCCAACTGATGTAAGGAGCGGAC 303
 Db 82 GlyMetTyrHisVal1ThrAsnAspCysSerAsnSerSer1IleVal1yTyrG1yAlaAlaAsp 101
 QY 304 ATGATCATGCACACCCCGGGTGCCTGCGCTTCGCGAGAACAACTTCGCCGCTGC 363
 Db 102 Met1IleMetHisThrProG1yCysValProCysValArgG1yAsnAsnSerSerArgCys 121
 QY 364 TGGGTAGCCGCTACCCCGACGCTCGAGCTAGAGAACGCCGCTCCCACTAGACAAATA 423
 Db 122 TyrValAlaLeuThrProThrLeuAlaAlaArgAlaSerValProThrThrThr1Ile 141
 QY 424 CGAGCGACGATGATTTGCTGCTGGGGCGGCTTCTTGTTCGGTATGTAAGGAGCGG 483
 Db 142 ArgArgHisValAspLeuLeuValG1yAlaAlaAlaPheCysSerAlaMetTyrValG1y 161
 QY 484 GATCTGCGGATCTGTCTTCTGCTGCTCCAGCTTCACCACTTCGCGCTCCCGGAT 543
 Db 162 AspLeuCysG1ySerValPheLeuValSerGlnLeuPheThr1IleSerProAlaArgHis 181
 QY 544 GAAGACGGGTGAAGTACATGCAATGCTCAATATCCGCGACATACAGGCTACGCTATG 603
 Db 182 GluThrValGlnAspCysAsnCysSer1IleTyrProG1yHis1IleThrG1yHisArgMet 201
 QY 604 GCTTGGATGTGATGATGACATGG 627
 Db 202 AlaTyrAspMetMetCAsnTyr 209

 RESULT 8
 US-08-927-597-48
 ; Sequence 48, Application US/08927597
 ; Patent No. 6245503
 ; GENERAL INFORMATION:

APPLICANT: MAERTENS, GEERT
 APPLICANT: BOSMAN, FONS
 APPLICANT: DE MARTYNOFF, GUY
 APPLICANT: BUOSE, MARIE-ANGE
 TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
 TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
 NUMBER OF SEQUENCES: 111
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: NIXON & VANDERHUYE P.C.
 STREET: 1100 NORTH GLEBE ROAD
 CITY: ARLINGTON
 STATE: VIRGINIA
 COUNTRY: U.S.A.
 ZIP: 22201-4714
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/927,597
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/612,973
 FILING DATE: 11-MAR-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: BYRNE, THOMAS E.
 REGISTRATION NUMBER: 32,205
 REFERENCE/DOCKET NUMBER: 1487-10
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 816-4000
 TELEFAX: (703) 816-4100
 INFORMATION FOR SEQ ID NO: 48:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 692 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-927-597-48

 Alignment Scores:
 Pred. No.: 1.59e-109 Length: 692
 Score: 1108.00 Matches: 206
 Percent Similarity: 100.00% Conservative: 2
 Best Local Similarity: 99.04% Mismatches: 0
 Query Match: 95.93% Indels: 0
 DB: 4 Gaps: 0

 US-09-899-303A-7 (1-633) x US-08-927-597-48 (1-692)
 QY 4 TTGGGTAAGTCAATCGATACCTTACGTGGCGCTTCCGCGACCTCATGGGCTACATTCCG 63
 Db 2 LeuG1yLysVal11IleAspThrLeuThrCysG1yPheAlaAspLeuValG1yTyr1IlePro 21
 QY 64 CTCGTGGGCGCCCGCCAGGCGGGGCGTCCAGAGCCCTGGCGCATGGCGCGTTCGTTCTG 123
 Db 22 LeuValG1yAlaProLeuG1yG1yAlaAlaArgAlaLeuAlaHisG1yValAlaGValLeu 41
 QY 124 GAAGACGGCGGTGAATGACAGACAGGGAATTTGCGTGGTTCCTCTTCTATCTTCCTC 183
 Db 42 GluAspG1yAlaAsnTyrAlaThrG1yAsnLeuProG1yLysSerPheSer1IlePheLeu 61
 QY 184 TTGGCTTTACTGCTCCGTGACCAATCCAGCTTCGCTTATGAGTGCGCCCAACGCTGCC 243
 Db 62 LeuAlaLeuLeuSerCysLeuThrValProAlaSerAlaTyrG1yValAlaGAsnValSer 81
 QY 244 GGGATGTACCATGTCAAGACGACTGCTCCAACTGATGTAAGGAGCGGAC 303
 Db 82 GlyMetTyrHisVal1ThrAsnAspCysSerAsnSerSer1IleVal1yTyrG1yAlaAlaAsp 101
 QY 304 ATGATCATGCACACCCCGGGTGCCTGCGCTTCGCGAGAACAACTTCGCCGCTGC 363
 Db 102 Met1IleMetHisThrProG1yCysValProCysValArgG1yAsnAsnSerSerArgCys 121

Db 102 MetIleMethIstHrProGlyCysValProCysValArgGluAsnSerSerArgCys 121
QY 364 TGGGTACCGCTCACCCCGCTGAGTACAGAACCCAGCTCCCGCTACGACATA 423
Db 122 TrpValAlaLeuThrProThrLeuAlaAlaArgAsnAlaSerValProThrThrIle 141
QY 424 CGAGCCAGCTGCATTGCTGCTGGGGCGGCTTCTTCTCCGTATGACGTGGG 483
Db 142 ArgArgHisValAspLeuLeuValGlyAlaAlaAlaPheCysSerAlaMetLysValGly 161
QY 484 GATCTCGCGGATCTCTCTCCGCTCCAGCTGTTACACATCTGCGCGCGCGCAT 543
Db 162 AspLeuGlySerValPheLeuValSerGlnLeuPheThrIleSerProAlaArgHis 181
QY 544 GAGACGCTGACAGACTCAATTGCTCAATCTATCCCGCCACATACAGCTACCGTATG 603
Db 182 GluThrValGlnAspCysAsnCysSerIleLysProGlyHisIleThrGlyHisArgMet 201
QY 604 GCTTGGATATGATGATGAACCTGC 627
Db 202 AlaTrpAspMetMetMetAsnTrp 209

RESULT 9

US-08-612-973-50
; Sequence 50, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHAYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-Mar-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 809 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-612-973-50

Alignment Scores:
Pred. No.: 1.69e-109
Score: 1108.00
Percent Similarity: 100.00%
Best Local Similarity: 99.04%
Query Match: 95.93%
DB: 4
Length: 809
Matches: 206
Conservative: 2
Mismatch: 0
Indels: 0
Gaps: 0

US-09-899-303a-7 (1-633) x US-08-612-973-50 (1-809)

QY 4 TTGGGTAAAGGTCATCATACCTTACGTCGGGCTTCCCGACCTCATGGGTATCTCCG 63
Db 119 LeuGlyLysValIleAspThrLeuThrCysGlyPheAlaAspLeuValGlyTyrIlePro 138
QY 64 CTCGTGCGCGCGCGCTAGGGGGTGTGCCAGAGCCCTGGCGGCAATGGCGTCCGTTCTG 123
Db 139 LeuValGlyAlaProLeuGlyGlyAlaAlaArgAlaLeuAlaHisGlyValArgValLeu 158
QY 124 GAAGACCGCGCTGACACTATGCAACAGGGAATTTGCTGCTGCTTCTTCTATCTTCTC 183
Db 159 GluAspIleValAsnTyrAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeu 178
QY 184 TTGCTTACTGTCTCTGTACACCTTCAGCTTCCAGCTTCCGCTATGAGTGCGCAACGTGTC 243
Db 179 LeuAlaLeuLeuSerCysLeuThrValProAlaSerAlaTyrGluValArgAsnValSer 198
QY 244 GCGATGTACCATGTACAGAACGACTGTCCCAACTCAAGCATTTGTATGAGCGACGCGAC 303
Db 199 GlyMetLysHisValThrAsnAspCysSerAsnSerSerIleValTyrGluAlaAlaAsp 218
QY 304 ATGATCATGCACACCCCGGGGTGCGTCCCTGCGTGGGAGAAACAATCTTCCGCTGC 363
Db 219 MetIleMethIstHrProGlyCysValProCysValArgGluAsnAsnSerSerArgCys 238
QY 364 TGGGTACCGCTCACCCCGCTGAGTACAGAACCCAGCGTCCCGCTACGACATA 423
Db 239 TrpValAlaLeuThrProThrLeuAlaAlaArgAsnAlaSerValProThrThrIle 258
QY 424 CGAGCCAGCTGCATTGCTGCTGGGGCGGCTTCTTCTCCGTATGACGTGGG 483
Db 259 ArgArgHisValAspLeuLeuValGlyAlaAlaAlaPheCysSerAlaMetLysValGly 278
QY 484 GATCTCGCGGATCTCTCTCCGCTCCAGCTGTTACACATCTGCGCGCGCAT 543
Db 279 AspLeuGlySerValPheLeuValSerGlnLeuPheThrIleSerProAlaArgHis 298
QY 544 GAGACGCTGACAGACTCAATTGCTCAATCTATCCCGCCACATACAGCTACCGTATG 603
Db 299 GluThrValGlnAspCysAsnCysSerIleLysProGlyHisIleThrGlyHisArgMet 318
QY 604 GCTTGGATATGATGATGAACCTGC 627
Db 319 AlaTrpAspMetMetMetAsnTrp 326

RESULT 10

US-08-927-597-50
; Sequence 50, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHAYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597
; FILING DATE:

CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/612,973
 FILING DATE: 11-MAR-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: BYRNE, THOMAS E.
 REGISTRATION NUMBER: 32,205
 REFERENCE/DOCKET NUMBER: 1487-10
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 816-4000
 TELEFAX: (703) 816-4100
 INFORMATION FOR SEQ ID NO: 50:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 809 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-927-597-50

Alignment Scores:
 Pred. No.: 1 69e-109 Length: 809
 Score: 1108.00 Matches: 206
 Percent Similarity: 100.00% Conservative: 2
 Best Local Similarity: 99.04% Mismatches: 0
 Query Match: 95.93% Indels: 0
 DB: 4 Gaps: 0

US-09-899-303a-7 (1-633) x US-08-927-597-50 (1-809)

```

QY 4 TTGGGTAAAGTATCGATACCTTACGTCGGGCTTCGCCGACCTCATGGGGTACATTCG 63
DB 119 LeuGlyLysValIleAspHrLeuThrCysGlyPheAlaAspLeuValGlyTrillePro 138
QY 64 CTCGTGGGCCCCCTAGGGGGTCTGCCAGAGCCCTGGCGCATGGCGGTCTG 123
DB 139 LeuValGlyAlaProLeuValGlyAlaAlaArgAlaLeuAlaHisGlyValAlaArgValLeu 158
QY 124 GAAGACGGGGTGAATGACACAGGGAATTCGCTGCTTCCTATCTCCCTC 183
DB 159 GluAspGlyValAsnTrpAlaThrGlyAsnLeuProGlyCysSerHeserillePheLeu 178
QY 184 TTGGCTTACTGTCTGTCTGACCATTCAGCTTCGCTTATAGGTGGCAGCATGTC 243
DB 179 LeuAlaLeuLeuSerCysLeuThrValProAlaSerAlaTrgIuValAlaArgAsnValSer 198
QY 244 GGGATGACCATGTCAGCAACGACTGCTCCAACTCAAGCATTTGTATAGAGCGCGGAC 303
DB 199 GlyMetLysHisValThrAsnAspCysSerAsnSerSerilleValTrgIuAlaAlaAsp 218
QY 304 ATGATGATGACACACCCCGGGTGGCGCTGCGGCTCGGGAGAACAACTTCCCGCTGC 363
DB 219 MetIleMetHisThrProGlyCysValProCysValAlaTrgIuAsnAsnSerSerArgCys 238
QY 364 TTGGTACGCTCACCSCACGCTCGCAGCTAGGAGCCAGCGCTCCCACTACGACAAATA 423
DB 239 TrpValAlaLeuThrProThrLeuAlaAlaArgAsnAlaSerValProThrTrille 258
QY 424 CCAAGCCACGCTGATGCTGCTGGGGGGCTGCTTGTGTTCCGATAGTACGTTGG 483
DB 259 ArgArgHisValAspLeuLeuValGlyAlaAlaAlaPheCysSerAlaMetTrgValGly 278
QY 484 GATCTGTGGGATCTGTCTGCTGCTCCGCTTACCATCTCCGCTCGCGCGCAT 543
DB 279 AspLeuCysGlySerValPheLeuValSerGlnLeuPheThrIleSerProAlaArgHis 298
QY 544 GAGACGGTCAAGACTGCAATGCTCAATCTATCCGGGCACATAACAGGTACCGTATG 603
DB 299 GluThrValGlnAspCysAsnSerIleTrpProGlyHisIleThrGlyHisAlaTrgMet 318
QY 604 GCTTGAGATATGATGATGAACG 627
DB 319 AlaTrpAspMetMetMetAsnTrp 326

```

RESULT 11
 US-08-462-195-2

Sequence 2, Application US/08462195
 Patent No. 5789544
 GENERAL INFORMATION:
 APPLICANT: MIYAMURA, TATSUO
 APPLICANT: SATO, IZUMI
 APPLICANT: MATSURA, YOSHIMARU
 APPLICANT: HONDA, YOSHIKAZU
 APPLICANT: SEKI, MAKOTO
 TITLE OF INVENTION: METHOD FOR PRODUCING ECTOPROTEIN OF
 HEPATITIS C VIRUS
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: OBION, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
 ADDRESSEE: P.C.
 STREET: 1755 S. Jefferson Davis Highway, Suite 400
 CITY: Arlington
 STATE: Virginia
 COUNTRY: U.S.A.
 ZIP: 22202

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/462,195
 FILING DATE: 05-JUN-1995
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/446,303
 FILING DATE: 22-MAY-1995
 APPLICATION NUMBER: US 08/074,584
 FILING DATE: 11-JUN-1993
 APPLICATION NUMBER: JP 152487/1992
 FILING DATE: 11-JUN-1992

ATTORNEY/AGENT INFORMATION:
 NAME: Obion, No. 5789544man F.
 REGISTRATION NUMBER: 24,618
 REFERENCE/DOCKET NUMBER: 4169-003-0
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 413-3000
 TELEFAX: (703) 413-2220
 TELEX: 248855 OPAT UR

INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 340 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-462-195-2

Alignment Scores:
 Pred. No.: 4 24e-107 Length: 340
 Score: 1084.00 Matches: 199
 Percent Similarity: 98.56% Conservative: 6
 Best Local Similarity: 95.67% Mismatches: 3
 Query Match: 93.85% Indels: 0
 DB: 1 Gaps: 0

US-09-899-303a-7 (1-633) x US-08-462-195-2 (1-340)

```

QY 4 TTGGGTAAAGTATCGATACCTTACGTCGGGCTTCGCCGACCTCATGGGGTACATTCG 63
DB 119 LeuGlyLysValIleAspHrLeuThrCysGlyPheAlaAspLeuValGlyTrillePro 138
QY 64 CTCGTGGGCCCCCTAGGGGGTCTGCCAGAGCCCTGGCGCATGGCGGTCTG 123
DB 139 LeuValGlyAlaProLeuValGlyAlaAlaArgAlaLeuAlaHisGlyValAlaArgValLeu 158
QY 124 GAAGACGGGGTGAATGACACAGGGAATTCGCTGCTTCCTATCTCCCTC 183

```

```

Db 159 GIuAspGLyValAsnTYrAlaThrGLyAsnLeuProGLyCysSerPheSerIlePheLeu 178
OY 184 TTGGCTTACTGTCTGTCTGTCAGCAATTCAGCTCCGTTATAGATGGCAACGTCTCC 243
Db 179 LeuAlaLeuLeuSerCysLeuThrIleProAlaSerAlaTYrGLuValArgAsnValSer 198
OY 244 GGGATGTACCATGTACAGACAGACATGCTCCAACTCAACATGATTTGTGTATAGAGCAGCGAC 303
Db 199 GYIleTYrHisValThrAsnAspCysSerAsnSerIleValTYrGLuAlaAlaAsp 218
OY 304 ATGATCATGACACACCCCGGGGTGCTGCGCTGCTGGAGAACACATCTCCCGCTGC 363
Db 219 ValIleMetHisAlaProGLyCysValProCysValArgGLuAsnAsnSerIleProCys 238
OY 364 TGGGTAGCGCTACCCCGCAGCTGCTAGCAAGCCAGCGTCCCACTAGCAACAATA 423
Db 239 TrpValAlaLeuThrProThrLeuAlaAlaArgAsnAlaSerValProThrThrIleu 258
OY 424 CGAGCCGACGTGATTTCTGCTGGGGCGGTCTTCTGTCTGCTATGACGTGGGG 483
Db 259 ArgArgHisValAspLeuLeuValGLyThrAlaAlaPheCysSerAlaMetTYrValGLy 278
OY 484 GATCTCTGCGATCTGTCTCTGCTCCAGCTGTCCAGCTGTCCACATCTCGCTCGCGGAT 543
Db 279 AspLeuCysGLySerValPheLeuIleSerGLnLeuPheThrPheSerProArgArgHis 298
OY 544 GAGACGGTGCAGAGACTGCAATTGCTCAATCTATCCCGGCACATACAGATCACCGGTATG 603
Db 299 GluThrValGLnAspCysAsnCysSerIleTYrProGLyHisValSerGLyHisArgMet 318
OY 604 GCTTGGGATATGATGATGAACCTGG 627
Db 319 AlaTrpAspMetMetMetAsnTrp 326

RESULT 12
US-08-636-883-2
: Sequence 2, Application US/08636883
: Patent No. 5830691
: GENERAL INFORMATION:
: APPLICANT: MIYAMURA, TATSUO
: APPLICANT: SAITO, IZUMU
: APPLICANT: MATSURA, YOSHIMARU
: APPLICANT: HONDA, YOSHIKAZU
: APPLICANT: SEKI, MAKOTO
: TITLE OF INVENTION: METHOD FOR PRODUCING ECTOPROTEIN OF
: NUMBER OF SEQUENCES: 9
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
: ADDRESSEE: P.C.
: STREET: 1755 S. Jefferson Davis Highway, Suite 400
: CITY: Arlington
: STATE: Virginia
: COUNTRY: U.S.A.
: ZIP: 22202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/636,883
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/08/446,303
: FILING DATE:
: APPLICATION NUMBER: US 08/074,584
: FILING DATE: 11-JUN-1993
: APPLICATION NUMBER: JP 152487/1992
: FILING DATE: 11-JUN-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Oblon, No. 5830691man F.

```

```

: REGISTRATION NUMBER: 24,618
: REFERENCE/DOCKET NUMBER: 4169-003-0
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 413-3000
: TELEFAX: (703) 413-2220
: TELEX: 248855 OPAT UR
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 340 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-636-883-2

Alignment Scores:
Pred. No.: 4,24e-107 Length: 340
Score: 1084.00 Matches: 199
Percent Similarity: 98.56% Conservative: 6
Best Local Similarity: 95.67% Mismatches: 3
Query Match: 93.85% Indels: 0
DB: 2 Gaps: 0

US-09-899-303A-7 (1-633) x US-08-636-883-2 (1-340)

OY 4 TTGGGTAAAGTCAATGATACCTTACGTGCGGCTTCCGCGCACTTCATGGGTACATTCG 63
Db 119 LeuGLyValIleAspThrLeuThrCysGLyPheAlaAspLeuMetGLyTYrIlePro 138
OY 64 CTCGTGCGGCGGCGGCTTGGGGGTGCTCCAGAGCCCTGGGCGCATAGGGCTCGGGTTTG 123
Db 139 LeuValGLyAlaProLeuGLyGLyAlaAlaArgAlaLeuAlaHisGLyValArgValLeu 158
OY 124 GAACAGCGCGTGAACCTATGCAACAGAGGAATTTCCTGCTGCTCTTCTATGCTCTC 183
Db 159 GIuAspGLyValAsnTYrAlaThrGLyAsnLeuProGLyCysSerPheSerIlePheLeu 178
OY 184 TTGGCTTACTGTCTGTCTGTCAGCAATTCAGCTCCGTTATAGATGGCAACGTCTCC 243
Db 179 LeuAlaLeuLeuSerCysLeuThrIleProAlaSerAlaTYrGLuValArgAsnValSer 198
OY 244 GGGATGTACCATGTACAGACAGACTGCTCAACATCAACATGCTATAGAGCAGCGGAC 303
Db 199 GYIleTYrHisValThrAsnAspCysSerAsnSerIleValTYrGLuAlaAlaAsp 218
OY 304 ATGATCATGACACACCCCGGGGTGCTGCGCTGCTGGAGAACATCTCCCGCTGC 363
Db 219 ValIleMetHisAlaProGLyCysValProCysValArgGLuAsnAsnSerIleProCys 238
OY 364 TGGGTAGCGCTACCCCGCAGCTGCTAGCAAGCCAGCGTCCCACTAGCAACAATA 423
Db 239 TrpValAlaLeuThrProThrLeuAlaAlaArgAsnAlaSerValProThrThrIleu 258
OY 424 CGAGCCGACGTGATTTCTGCTGGGGCGGTCTTCTGTCTGCTATGACGTGGGG 483
Db 259 ArgArgHisValAspLeuLeuValGLyThrAlaAlaPheCysSerAlaMetTYrValGLy 278
OY 484 GATCTCTGCGATCTGTCTCTGCTCCAGCTGTCCAGCTGTCCACATCTCCGCGGAT 543
Db 279 AspLeuCysGLySerValPheLeuIleSerGLnLeuPheThrPheSerProArgArgHis 298
OY 544 GAGACGGTGCAGAGACTGCAATTGCTCAATCTATCCCGGCACATACAGATCACCGGTATG 603
Db 299 GluThrValGLnAspCysAsnCysSerIleTYrProGLyHisValSerGLyHisArgMet 318
OY 604 GCTTGGGATATGATGATGAACCTGG 627
Db 319 AlaTrpAspMetMetMetAsnTrp 326

RESULT 13
US-09-127-829-2
: Sequence 2, Application US/09127829
: Patent No. 6063904
: GENERAL INFORMATION:

```

APPLICANT: MIYAMURA, TATSUO
APPLICANT: SAITO, IZUMU
APPLICANT: MATSUDRA, YOSHINARU
APPLICANT: HONDA, YOSHIKAZU
APPLICANT: SEKI, MAKOTO
TITLE OF INVENTION: METHOD FOR PRODUCING ECTOPROTEIN OF
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/127, 829
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/446, 303
FILING DATE:
APPLICATION NUMBER: US 08/074, 584
FILING DATE: 11-JUN-1993
APPLICATION NUMBER: JP 152487/1992
FILING DATE: 11-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NO. 6063904man F.
REGISTRATION NUMBER: 24, 618
REFERENCE/DOCKET NUMBER: 4169-003-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 340 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-127-829-2
Alignment Scores:
Pred. No.: 4,24e-107 Length: 340
Score: 1084.00 Matches: 199
Percent Similarity: 98.56% Conservative: 6
Best Local Similarity: 95.67% Mismatches: 3
Query Match: 93.85% Indels: 0
Gaps: 0
DB: 3
US-09-899-303a-7 (1-633) x US-09-127-829-2 (1-340)
QY 4 TTGGGTAAGGTCAATGATACCTTACGTCGGGCTGCGGACCTCATGGGTCATTCGCG 63
DB 119 LGGGLyLVsVAlIleAspHrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIlePro 138
QY 64 CTGCTGGCGCCGCCCTAGGGGGTGTGCCAGAGCCCTGGCGCATGGCGTCCGGGTTCTG 123
DB 139 LGGGLyLVsVAlIleAspHrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIlePro 138
QY 124 GAAGAGGGCGCTGAACATGCAACAGGGAATTTGGCTGCTTCTTCTATCTTCCTC 133
DB 159 GLAAspGLyValAsnTyrAlaThrIleGlyAsnLeuProGlyCysSerPheSerIlePheLeu 178
QY 184 TTGGCTTATGTCCTGCTGACCATTCGACCTTCGCTATGAGGGGCGGCACAGTGTCC 243
DB 179 LGGGLyLVsVAlIleAspHrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIlePro 138

QY 244 GGGATGTACCATGTGTCAGACGACTGCTCCAACTCAAGCATTTGTATGAGCGGCGAC 303
DB 199 GLYIleTyrIleValThrAsnAspCysSerAsnSerIleValTyrGluAlaIleAsp 218
QY 304 ATGATCATGCAACACCCCGGGTGGTCCCTGCTTCGGGAGAACACTTTCGCCGTGC 363
DB 219 ValIleMetHisAlaProGlyCysValProCysValAlaGlyIleAsnSerSerArgCys 238
QY 364 TTGGTACCGCTCCGCCAGCTGCGACGTAGAGAGCCAGCGCTCCCACTAGACAATA 423
DB 239 TrpValAlaLeuThrProThrLeuAlaAlaArgAlaAsnAlaSerValProThrThrIleLeu 258
QY 424 CGAGCGCCAGCTGCATTTGCTGTTGGGGCGGCTTTGTTCCGTATGATGCGGGG 483
DB 259 ArgAlaThrIleValAsnLeuLeuValGlyThrAlaAlaPheCysSerIleAlaMetTyrValGly 278
QY 484 GATCTGCGCGGATCTGCTCTGCTGCTCCAGCTGTTCACCATTCGCGCTCGCGGAT 543
DB 279 AspLeuGlySerValPheLeuIleSerGlnLeuPheThrPheSerProAlaArgHis 298
QY 544 GAGAGGCTGACGAGCATGCTCATCTATCCGGCCACATACAGGTACACCGTATG 603
DB 299 GluThrValGlnAspCysAsnCysSerIleTyrProGlyHisValSerGlyHisArgMet 318
QY 604 GCTTGGGATATGATGATGAACGTG 627
DB 319 AlaTrpAspMetMetIleAsnTrp 326
RESULT 14
US-09-014-416-3
Sequence 3, Application US/09014416
Patent No. 6153421
GENERAL INFORMATION:
APPLICANT: Yanagi, Masayuki
APPLICANT: Buhk, Jens
APPLICANT: Emerson, Susanne U.
APPLICANT: Purcell, Robert H.
TITLE OF INVENTION: CLONED GENOMES OF INFECTIOUS HEPATITIS C VIRUSES AND
TITLE OF INVENTION: USES THEREOF
FILE REFERENCE: 20264276
CURRENT APPLICATION NUMBER: US/09/014, 416
CURRENT FILING DATE: 1998-01-27
EARLIER APPLICATION NUMBER: US 60/053, 062
EARLIER FILING DATE: 1997-07-18
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 3010
TYPE: PRT
ORGANISM: Hepatitis C virus
US-09-014-416-3
Alignment Scores:
Pred. No.: 2,17e-106 Length: 3010
Score: 1081.00 Matches: 199
Percent Similarity: 98.56% Conservative: 6
Best Local Similarity: 95.67% Mismatches: 3
Query Match: 93.59% Indels: 0
Gaps: 0
DB: 4
US-09-899-303a-7 (1-633) x US-09-014-416-3 (1-3010)
QY 4 TTGGGTAAGGTCAATGATACCTTACGTCGGGCTGCGGACCTCATGGGTCATTCGCG 63
DB 119 LGGGLyLVsVAlIleAspHrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIlePro 138
QY 64 CTGCTGGCGCCGCCCTAGGGGGTGTGCCAGAGCCCTGGCGCATGGCGTCCGGGTTCTG 123
DB 139 LGGGLyLVsVAlIleAspHrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIlePro 138
QY 124 GAAGAGGGCGCTGAACATGCAACAGGGAATTTGGCTGCTTCTTCTATCTTCCTC 183

```

Db 159 GluaspelYalaSnYrAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeu 178
QY 184 TTGGCTTTACTGTCCTGTCTGACCAATTCACGCTTCGCTTAATGAGTGGCCAGCACTGTC 243
Db 179 LeuAlaLeuLeuSerCysLeuThrIleProAlaSerAlaTyrgIuValaArgAsnValSer 198
QY 244 GGGATGTACCATGTGCAGAGAGAGACTGCTCCAACTCAAGCATGTGTATGAGGCGGCGAC 303
Db 199 GlyIleTyrlAsValThrAsnAspCysSerAsnSerIleValTyrgIuAlaAlaAsp 218
QY 304 ATGATCATGACACACCCCGGGTGGCGCCCGCTGCGTGGGAGAACACTGTTCGGCTGC 363
Db 219 ValIleMetHisThrProGlyCysValProCysValGlnGluGlyAsnSerSerArgCys 238
QY 364 TGGGTACCGCTCACCCCGCTGCGCTGAGTAAGAACGCCAGCGTCCCACTACGACATA 423
Db 239 TrpValAlaLeuThrProThrIleuAlaAlaArgsnAlaSerValProThrThrIle 258
QY 424 CGAGCCACGTCGATTTGCTGCTGGGGCGGCTCTTGTCTCCGTATGTACGTGGG 483
Db 259 ArgArgHisValAspLeuLeuValGlyThrAlaAlaPheCysSerAlaMetTyValGly 278
QY 484 GATCTCGCGGATGTCCTGCTCCGCTCCAGCTGTTCACCATGTGCGCGCGGCGAT 543
Db 279 AspLeuCysGlySerIlePheLeuValSerGlnLeuPheThrPheSerProAlaArgHis 298
QY 544 GAGACGGTGCAGACTCAATTTGCTCAATCATCCCGGCCACATACAGAGTCAACGCTATG 603
Db 299 GluThrValGlnAspCysAsnCysSerIleTyrgProGlyHisValSerGlyHisArgMet 318
QY 604 GCTTGGATATGATGATGAAGTGG 627
Db 319 AlaTrpAspMetMetMetAsnTrp 326

```

RESULT 15

```

US-09-014-416-5
; Sequence 5, Application US/09014416
; Patent No. 6153421
; GENERAL INFORMATION:
; APPLICANT: Yanagi, Masayuki
; APPLICANT: Bukh, Jens
; APPLICANT: Emerson, Susanne U.
; APPLICANT: Purcell, Robert H.
; TITLE OF INVENTION: CLONED GENOMES OF INFECTIOUS HEPATITIS C VIRUSES AND
; FILE REFERENCE: 20264276
; CURRENT APPLICATION NUMBER: US/09/014,416
; EARLIER FILING DATE: 1998-01-27
; EARLIER APPLICATION NUMBER: US 60/053,062
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 3011
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-014-416-5

```

Alignment Scores:

```

Pred. No.: 2,17e-106 Length: 3011
Score: 1081.00 Matches: 199
Percent Similarity: 98.56% Conservative: 6
Best Local Similarity: 95.67% Mismatches: 3
Query Match: 93.59% Indels: 0
DB: 4 Gaps: 0

```

US-09-899-303a-7 (1-633) x US-09-014-416-5 (1-3011)

```

QY 4 TTGGGTAAAGTCATGATACCTTACGTCGCGCTTCGCCGACCTACATGAGGCTACATTCGG 63
Db 119 LeuclYlyValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrlIlePro 138
QY 64 CTCGTGGCGCCCGCCCTAGGGGGGTGTGCCAGAGCCCTGGCGCATGGCGGTCCGGGTTCTG 123

```

```

Db 139 LeuValGlyAlaProLeuGlyGlyAlaAlaAlaArgAlaLeuAlaHisGlyValaArgValLeu 158
QY 124 GAAGACGGCGCTGAACATAATGCAACAGGAATTTGCTGCTGCTCTTCTATCTTCTCTC 183
Db 159 GluaspelYalaSnYrAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeu 178
QY 184 TTGGCTTTACTGTCCTGTCTGACCAATTCACGCTTCGCTTAATGAGTGGCCAGCACTGTC 243
Db 179 LeuAlaLeuLeuSerCysLeuThrIleProAlaSerAlaTyrgIuValaArgAsnValSer 198
QY 244 GGGATGTACCATGTGCAGAGAGAGACTGCTCCAACTCAAGCATGTGTATGAGGCGGCGAC 303
Db 199 GlyIleTyrlAsValThrAsnAspCysSerAsnSerIleValTyrgIuAlaAlaAsp 218
QY 304 ATGATCATGACACACCCCGGGTGGCGCCCGCTGCGTGGGAGAACACTGTTCGGCTGC 363
Db 219 ValIleMetHisThrProGlyCysValProCysValGlnGluGlyAsnSerSerArgCys 238
QY 364 TGGGTACCGCTCACCCCGCTGCGCTGAGTAAGAACGCCAGCGTCCCACTACGACATA 423
Db 239 TrpValAlaLeuThrProThrIleuAlaAlaArgsnAlaSerValProThrThrIle 258
QY 424 CGAGCCACGTCGATTTGCTGCTGGGGCGGCTCTTGTCTCCGTATGTACGTGGG 483
Db 259 ArgArgHisValAspLeuLeuValGlyThrAlaAlaPheCysSerAlaMetTyValGly 278
QY 484 GATCTCGCGGATGTCCTGCTCCGCTCCAGCTGTTCACCATGTGCGCGCGGCGAT 543
Db 279 AspLeuCysGlySerIlePheLeuValSerGlnLeuPheThrPheSerProAlaArgHis 298
QY 544 GAGACGGTGCAGACTCAATTTGCTCAATCATCCCGGCCACATACAGAGTCAACGCTATG 603
Db 299 GluThrValGlnAspCysAsnCysSerIleTyrgProGlyHisValSerGlyHisArgMet 318
QY 604 GCTTGGATATGATGATGAAGTGG 627
Db 319 AlaTrpAspMetMetMetAsnTrp 326

```

Search completed: February 12, 2003, 08:56:18
Job time : 21 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 18, 2003, 04:52:49 ; Search time 220 Seconds
(without alignments)
6479.613 Million cell updates/sec

Title: US-09-899-303a-7

Perfect score: 633
Sequence: 1 ATGTTGGGTAAAGTCATCGA.....TGATGATCACTGGTAATAG 633

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

N.Geneseq_101002.*
1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	633	100.0	633	17	AAT12706
2	612.8	96.8	636	17	AAT12709
3	608.8	96.2	795	17	AAT12705
4	605.8	95.7	2086	17	AAT12973
5	605.8	95.7	2433	17	AAT12974
6	564.8	89.2	673	19	AAV42305
7	564.6	89.2	2187	19	ABA03491
8	564.6	89.2	2540	14	AA043889
9	564.6	89.2	2540	15	AA063753

10	564.6	89.2	9605	24	AAD25332
11	564.6	89.2	11062	24	AAD25331
12	564.6	89.2	11076	21	AAA98965
13	563	88.9	1880	13	AA024467
14	561.4	88.7	932	13	AA020923
15	561.4	88.7	2540	13	AA029628
16	559.8	88.4	1863	12	AAQ15363
17	558.2	88.2	1562	19	AAV60672
18	558.2	88.2	2829	19	AAV60673
19	557.8	88.1	1270	19	AAV60668
20	557.8	88.1	3461	15	AA064068
21	557.8	88.1	3461	15	AAT30386
22	556.2	87.9	1037	15	AA058449
23	553.4	87.4	1251	13	AA026981
24	553.4	87.4	2116	12	AAQ12242
25	553.4	87.4	2187	19	ABA03492
26	553.4	87.4	3360	17	AAT03677
27	553.4	87.4	9413	16	AAT03960
28	553.4	87.4	9413	16	AA081559
29	553.4	87.4	9413	16	AA080498
30	553.4	87.4	9413	24	AAD25517
31	553.4	87.4	9595	20	AAX24843
32	553.4	87.4	9595	22	AAC86939
33	553.4	87.4	9595	22	AAE23492
34	553.4	87.4	9599	20	AAX24833
35	550.8	87.0	1251	13	AAQ22610
36	550.8	87.0	1398	13	AAQ22140
37	550.2	86.9	1682	13	AAQ27159
38	549.8	86.9	3401	15	AA064069
39	549.8	86.9	3401	16	AAT30387
40	549.8	86.9	9609	24	AAD33038
41	549.6	86.8	9436	15	AA063499
42	548.2	86.6	1734	14	AA040437
43	548.2	86.6	1734	14	AA040438
44	546.2	86.3	964	13	AAQ25745
45	543.4	85.8	1734	14	AAQ40428

ALIGNMENTS

RESULT 1
AAT12706
ID AAT12706 standard; DNA: 633 BP.
XX
AC AAT12706;
XX
DT 23-SEP-1996 (first entry)
XX
DE HCV E1 construct HCC111A.
XX
KW HCV E1; E2: disulphide bond cleavage; envelope protein; vaccine; human; serotype; reversed phase hybridisation assay; genotype; antigen; sera; ss.
XX
OS Hepatitis C virus.
XX
PN MO9604385-A2.
XX
PD 15-FEB-1996.
XX
PF 31-JUL-1995; 95WO-EP03031.
XX
PR 29-JUL-1994; 94EP-0870132.
XX
PA (INNO-) INNOGENETICS NV.
XX
PI Bosman F, Buysse M, De Martynoff G, Maertens G;
XX
DR WPI: 1996-129401/13.
XX
PT Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope proteins - In presence of disulphide bond cleavage agent, to

PT produce proteins suitable for direct use in vaccines or diagnostic assays of HCV

XX Claim 23; Fig 21; 146pp; English.

XX PS
XX AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1 and E2 protein coding sequence constructs. These sequences are included in vectors for the production of recombinant E1, E2, and E1/E2 proteins. CC
CC The recombinant proteins can then be isolated using a method of the invention. In the method, the envelope proteins are purified by CC
CC carrying out a disulphide bond cleavage, or a reduction step with a disulphide bond cleavage agent, after lysis of recombinant host cells. CC
CC The constructs containing the purified HCV envelope proteins can be used for vaccinating humans against HCV, for in vitro detection of HCV CC
CC antibodies in a sample, and in a serotyping assay for detecting one or more serological types of HCV present in a biological sample. The CC
CC constructs can also be immobilised on a solid substrate and incorporated into a reversed phase hybridisation assay for determining the presence or the genotype of HCV. The new purification method preserves the CC
CC conformation of the recombinantly expressed E1, E2 and E1/E2, and CC
CC eliminates contaminating proteins. Antigens isolated using this method are more reactive with human sera than those isolated by known CC
CC techniques.

SQ Sequence 633 BP; 111 A; 192 C; 174 G; 156 T; 0 other;

Query Match 100.0%; Score 633; DB 17; Length 633;

Best Local Similarity 100.0%; Pred. No. 5,7e-166;

Matches 633; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 ATGTGGGTAAAGTATGATACCTTACGTCGCGCTTCGCCGACCTATGGGGTACTT 60
DB 1 ATGTGGGTAAAGTATGATACCTTACGTCGCGCTTCGCCGACCTATGGGGTACTT 60
OY 61 CCGCTCTGCGGCGCCCTTAGGGGTCTGCAGAGCCCTGCGCCATGAGCGCTCCGGGTT 120
DB 61 CCGCTCTGCGGCGCCCTTAGGGGTCTGCAGAGCCCTGCGCCATGAGCGCTCCGGGTT 120
OY 61 CCGCTCTGCGGCGCCCTTAGGGGTCTGCAGAGCCCTGCGCCATGAGCGCTCCGGGTT 120
DB 61 CCGCTCTGCGGCGCCCTTAGGGGTCTGCAGAGCCCTGCGCCATGAGCGCTCCGGGTT 120
OY 121 CTGGAAGACGGCGTGAATATGCAACAGGGAATTTGCTGTGCTTCTATATCTTC 180
DB 121 CTGGAAGACGGCGTGAATATGCAACAGGGAATTTGCTGTGCTTCTATATCTTC 180
OY 121 CTGGAAGACGGCGTGAATATGCAACAGGGAATTTGCTGTGCTTCTATATCTTC 180
DB 121 CTGGAAGACGGCGTGAATATGCAACAGGGAATTTGCTGTGCTTCTATATCTTC 180
OY 181 CTCTGGTCTTACTCTCTGCTGTCATTCAGCTTCGCGTTAAGAGTGGCAACGTG 240
DB 181 CTCTGGTCTTACTCTCTGCTGTCATTCAGCTTCGCGTTAAGAGTGGCAACGTG 240
OY 181 CTCTGGTCTTACTCTCTGCTGTCATTCAGCTTCGCGTTAAGAGTGGCAACGTG 240
DB 181 CTCTGGTCTTACTCTCTGCTGTCATTCAGCTTCGCGTTAAGAGTGGCAACGTG 240
OY 241 TCCGGGATGACATGATGCAAGAACAGCTGTCACACTCAAGATGTGTATGAGCAGC 300
DB 241 TCCGGGATGACATGATGCAAGAACAGCTGTCACACTCAAGATGTGTATGAGCAGC 300
OY 241 TCCGGGATGACATGATGCAAGAACAGCTGTCACACTCAAGATGTGTATGAGCAGC 300
DB 241 TCCGGGATGACATGATGCAAGAACAGCTGTCACACTCAAGATGTGTATGAGCAGC 300
OY 301 GACATGATGATGACACACCCCGGGTGGCTGCCCTGCTTCCGGGAGAACACTTCCCGC 360
DB 301 GACATGATGATGACACACCCCGGGTGGCTGCCCTGCTTCCGGGAGAACACTTCCCGC 360
OY 301 GACATGATGATGACACACCCCGGGTGGCTGCCCTGCTTCCGGGAGAACACTTCCCGC 360
DB 301 GACATGATGATGACACACCCCGGGTGGCTGCCCTGCTTCCGGGAGAACACTTCCCGC 360
OY 361 TGTCTGGTAAAGGCTGACCCCGAGCTGCAGTAAAGACGCCAGGCTCCCACTACGACA 420
DB 361 TGTCTGGTAAAGGCTGACCCCGAGCTGCAGTAAAGACGCCAGGCTCCCACTACGACA 420
OY 361 TGTCTGGTAAAGGCTGACCCCGAGCTGCAGTAAAGACGCCAGGCTCCCACTACGACA 420
DB 361 TGTCTGGTAAAGGCTGACCCCGAGCTGCAGTAAAGACGCCAGGCTCCCACTACGACA 420
OY 421 ATACACACCCAGCTGATTTGCTTGGGGGCGTGTCTTCTGTCGCTATGACGTG 480
DB 421 ATACACACCCAGCTGATTTGCTTGGGGGCGTGTCTTCTGTCGCTATGACGTG 480
OY 421 ATACACACCCAGCTGATTTGCTTGGGGGCGTGTCTTCTGTCGCTATGACGTG 480
DB 421 ATACACACCCAGCTGATTTGCTTGGGGGCGTGTCTTCTGTCGCTATGACGTG 480
OY 481 GGGGATCTCTGGGATCTGCTTCTCTGCTGTCACAGCTGTCACCATCTCCGCTCCGG 540
DB 481 GGGGATCTCTGGGATCTGCTTCTCTGCTGTCACAGCTGTCACCATCTCCGCTCCGG 540
OY 481 GGGGATCTCTGGGATCTGCTTCTCTGCTGTCACAGCTGTCACCATCTCCGCTCCGG 540
DB 481 GGGGATCTCTGGGATCTGCTTCTCTGCTGTCACAGCTGTCACCATCTCCGCTCCGG 540
OY 541 CATGAGAGGTGACAGATGCAATGCTCAATCTATCCCGGACATATACAGGTACCGT 600
DB 541 CATGAGAGGTGACAGATGCAATGCTCAATCTATCCCGGACATATACAGGTACCGT 600
OY 541 CATGAGAGGTGACAGATGCAATGCTCAATCTATCCCGGACATATACAGGTACCGT 600
DB 541 CATGAGAGGTGACAGATGCAATGCTCAATCTATCCCGGACATATACAGGTACCGT 600
OY 601 ATGGCTTGGGATGATGATGAACTGTAATAG 633
DB 601 ATGGCTTGGGATGATGATGAACTGTAATAG 633
OY 601 ATGGCTTGGGATGATGATGAACTGTAATAG 633
DB 601 ATGGCTTGGGATGATGATGAACTGTAATAG 633

```

RESULT 2

ID AAT12709

XX AAT12709 standard; DNA; 636 BP.

XX AAT12709;

XX 23-SEP-1996 (first entry)

XX HCV E1 construct HCC117A.

XX HCV; E1, E2; disulphide bond cleavage; envelope protein; vaccine; human; serotype; reversed phase hybridisation assay; genotype; antigen; sera;

XX Hepatitis C virus.

XX W09604385-A2.

XX 15-FEB-1996.

XX 31-JUL-1995; 95WO-EP03031.

XX 29-JUL-1994; 94EP-0870132.

XX (INNO-) INNOGENETICS NV.

XX Bosman F, Buyse M, De Martynoff G, Maertens G;

XX WPI: 1996-129401/13.

XX purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope

XX proteins - in presence of disulphide bond cleavage agent, to

XX produce proteins suitable for direct use in vaccines or diagnostic

XX assays of HCV

PS Claim 23; Fig 21; 146pp; English.

XX AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1 and E2 protein coding sequence constructs. These sequences are included in vectors for the production of recombinant E1, E2, and E1/E2 proteins. CC
CC The recombinant proteins can then be isolated using a method of the invention. In the method, the envelope proteins are purified by CC
CC carrying out a disulphide bond cleavage, or a reduction step with a disulphide bond cleavage agent, after lysis of recombinant host cells. CC
CC The constructs containing the purified HCV envelope proteins can be used for vaccinating humans against HCV, for in vitro detection of HCV CC
CC antibodies in a sample, and in a serotyping assay for detecting one or more serological types of HCV present in a biological sample. The CC
CC constructs can also be immobilised on a solid substrate and incorporated into a reversed phase hybridisation assay for determining the presence or the genotype of HCV. The new purification method preserves the CC
CC conformation of the recombinantly expressed E1, E2 and E1/E2, and CC
CC eliminates contaminating proteins. Antigens isolated using this method are more reactive with human sera than those isolated by known CC
CC techniques.

SQ Sequence 636 BP; 111 A; 197 C; 175 G; 153 T; 0 other;

Query Match 96.8%; Score 612.8; DB 17; Length 636;

Best Local Similarity 96.1%; Pred. No. 2.3e-160;

Matches 620; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```

OY 1 ATGTGGGTAAAGTATGATACCTTACGTCGCGCTTCGCCGACCTATGGGGTACTT 60
DB 1 ATGTGGGTAAAGTATGATACCTTACGTCGCGCTTCGCCGACCTATGGGGTACTT 60
OY 61 CCGCTCTGCGGCGCCCTTAGGGGTCTGCAGAGCCCTGCGCCATGAGCGCTCCGGGTT 120
DB 61 CCGCTCTGCGGCGCCCTTAGGGGTCTGCAGAGCCCTGCGCCATGAGCGCTCCGGGTT 120
OY 61 CCGCTCTGCGGCGCCCTTAGGGGTCTGCAGAGCCCTGCGCCATGAGCGCTCCGGGTT 120
DB 61 CCGCTCTGCGGCGCCCTTAGGGGTCTGCAGAGCCCTGCGCCATGAGCGCTCCGGGTT 120
OY 121 CTGGAAGACGGCGTGAATATGCAACAGGGAATTTGCTGTGCTTCTATATCTTC 180
DB 121 CTGGAAGACGGCGTGAATATGCAACAGGGAATTTGCTGTGCTTCTATATCTTC 180

```

```

Db 121 CTGGAGAGAGGGGCTGAACATGACAAAGGAATTGCGGTGCTCTTCTTACTTC 180
Oy 181 CTTTGGCTTACTGCTGCTGACATTCAGCTTCCGCTTATGAGGCGCAACGTG 240
Db 181 CTTTGGCTTACTGCTGCTGACATTCAGCTTCCGCTTATGAGGCGCAACGTG 240
Oy 241 TCCGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
Db 241 TCCGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
Oy 301 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
Db 301 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
Oy 361 TCGTGGGTAGGCGCTACCCCGACGCTCGGCGTACGAGACGCCACATCCCTAC 420
Db 361 TCGTGGGTAGGCGCTACCCCGACGCTCGGCGTACGAGACGCCACATCCCTAC 420
Oy 421 ATACGACGCCACGTCGATTTGCTGTTGGGGGCGCTGCTTCTGTTCCGCTATG 480
Db 421 ATACGACGCCACGTCGATTTGCTGTTGGGGGCGCTGCTTCTGTTCCGCTATG 480
Oy 481 GGGGATCTTGGGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 540
Db 481 GGGGATCTTGGGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 540
Oy 541 CATGAGACGCTGACGACGACGACGACGACGACGACGACGACGACGACGACG 600
Db 541 CATGAGACGCTGACGACGACGACGACGACGACGACGACGACGACGACGACG 600
Oy 601 ATGCGTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATG 632
Db 601 ATGCGTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATG 632

```

RESULT 3

AAT12705 standard; DNA; 795 BP.

AAT12705;

23-SEP-1996 (first entry)

HCV E1 construct HCC110A.

HCV; E1; E2: disulphide bond cleavage; envelope protein; vaccine; human;

serotype; reversed phase hybridisation assay; genotype; antigen; sera;

89.

Hepatitis C virus.

W09604385-A2.

15-FEB-1996.

31-JUL-1995; 95WO-EP03031.

29-JUL-1994; 94EP-0870132.

(INNO-) INNOGENETICS NV.

Bosman F, Buyse M, De Martynoff G, Maertens G;

WPI, 1996-129401/13.

Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope proteins - in presence of disulphide bond cleavage agent, to produce proteins suitable for direct use in vaccines or diagnostic assays of HCV

Claim 23; Fig 21; 146pp; English.

AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1 and E2 protein coding sequence constructs. These sequences are included in vectors for the production of recombinant E1, E2, and E1/E2 proteins. The recombinant proteins can then be isolated using a method of the invention. In the method, the envelope proteins are purified by carrying out a disulphide bond cleavage, or a reduction step with a disulphide bond cleavage agent, after lysis of recombinant host cells. The constructs containing the purified HCV envelope proteins can be used for vaccinating humans against HCV, for in vitro detection of HCV antibodies in a sample, and in a serotyping assay for detecting one or more serological types of HCV present in a biological sample. The constructs can also be immobilised on a solid substrate and incorporated into a reversed phase hybridisation assay for determining the presence or the genotype of HCV. The new purification method preserves the conformation of the recombinantly expressed E1, E2 and E1/E2, and eliminates contaminating proteins. Antigens isolated using this method are more reactive with human sera than those isolated by known techniques.

Sequence 795 BP; 130 A; 240 C; 231 G; 194 T; 0 other;

Query Match 96.2%; Score 608.8; DB 17; Length 795;
Best Local Similarity 98.1%; Pred. No. 3.3e-159;
Matches 616; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```

Oy 1 ATGTTGGGTAAAGTATCATGATACCTTACGTGCGGCTTCCGACCTCATGGGATCAT 60
Db 1 ATGTTGGGTAAAGTATCATGATACCTTACGTGCGGCTTCCGACCTCATGGGATCAT 60
Oy 61 CCGCTGCTGGGCGCCCGCTAGGGGCTGTCGACAGCCCTGCGCAGTGGCTCCGGGTT 120
Db 61 CCGCTGCTGGGCGCCCGCTAGGGGCTGTCGACAGCCCTGCGCAGTGGCTCCGGGTT 120
Oy 121 CTGGAAGACGGCGGTGAACATGACAGAGGAATTTGCCGTTGCTTCTATCTTC 180
Db 121 CTGGAAGACGGCGGTGAACATGACAGAGGAATTTGCCGTTGCTTCTATCTTC 180
Oy 181 CTTTGGCTTACTGCTGCTGTCGACATTCAGCTTCCGCTTATGAGGCGCAACGTG 240
Db 181 CTTTGGCTTACTGCTGCTGTCGACATTCAGCTTCCGCTTATGAGGCGCAACGTG 240
Oy 241 TCCGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
Db 241 TCCGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
Oy 301 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
Db 301 GACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
Oy 361 TCGTGGGTAGGCGCTACCCCGACGCTCGGCGTACGAGACGCCACATCCCTAC 420
Db 361 TCGTGGGTAGGCGCTACCCCGACGCTCGGCGTACGAGACGCCACATCCCTAC 420
Oy 421 ATACGACGCCACGTCGATTTGCTGTTGGGGGCGCTGCTTCTGTTCCGCTATG 480
Db 421 ATACGACGCCACGTCGATTTGCTGTTGGGGGCGCTGCTTCTGTTCCGCTATG 480
Oy 481 GGGGATCTTGGGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 540
Db 481 GGGGATCTTGGGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 540
Oy 541 CATGAGACGCTGACGACGACGACGACGACGACGACGACGACGACGACGACG 600
Db 541 CATGAGACGCTGACGACGACGACGACGACGACGACGACGACGACGACGACG 600
Oy 601 ATGCGTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATG 628
Db 601 ATGCGTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATG 628

```

RESULT 4

AAT12973 standard; DNA; 2086 BP.

XX AC AAT12973;
 XX AC 24-SEP-1996 (first entry)
 XX DE HCV E1 construct HCC165.
 XX KW HCV; E1, E2; disulphide bond cleavage; envelope protein; vaccine; human;
 KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
 KW ss.
 XX OS Hepatitis C virus.
 XX OS W09604385-A2.
 XX PD 15-FEB-1996.
 XX PF 31-JUL-1995; 95WO-EP03031.
 XX PR 29-JUL-1994; 94EP-0870132.
 XX PA (INNO-) INNOGENETICS NV.
 XX PI Bosman F, Buyse M, De Martynoff G, Maertens G;
 XX DR WPI: 1996-129401/13.
 XX PT Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
 PT proteins - in presence of disulphide bond cleavage agent, to
 PT produce proteins suitable for direct use in vaccines or diagnostic
 PT assays of HCV
 XX
 XX PS Claim 23; Fig 21; 146pp; English.
 XX SS
 CC AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
 CC and E2 protein coding sequence constructs. These sequences are included
 CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
 CC The recombinant proteins can then be isolated using a method of the
 CC invention. In the method, the envelope proteins are purified by
 CC carrying out a disulphide bond cleavage, or a reduction step with a
 CC disulphide bond cleavage agent, after lysis of recombinant host cells.
 CC The constructs containing the purified HCV envelope proteins can be used
 CC for vaccinating humans against HCV, for in vitro detection of HCV
 CC antibodies in a sample, and in a serotyping assay for detecting one or
 CC more serological types of HCV present in a biological sample. The
 CC constructs can also be immobilised on a solid substrate and incorporated
 CC into a reversed phase hybridisation assay for determining the presence or
 CC the genotype of HCV. The new purification method preserves the
 CC conformation of the recombinantly expressed E1, E2 and E1/E2, and
 CC eliminates contaminating proteins. Antigens isolated using this method
 CC are more reactive with human sera than those isolated by known
 CC techniques.
 CC
 XX SQ Sequence 2086 BP; 366 A; 635 C; 601 G; 484 T; 0 other;
 Query Match 95.7%; Score 605.8; DB 17; Length 2086;
 Best Local Similarity 98.1%; Pred. No. 3.1e-158;
 Matches 613; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
 OY 4 TTGGTAAAGTCATGATGACCTTACGTCGCGGCTTCGCGACCTGATGGGTACTTCG 63
 DB 4 TTGGTAAAGTCATGATGACCTTACGTCGCGGCTTCGCGACCTGATGGGTACTTCG 63
 OY 64 CTCGTGCGCGCCCTTACGAGGCGTCCAGAGCCCTGCGCATGCGCTCGGGTTCTG 123
 DB 64 CTCGTGCGCGCCCTTACGAGGCGTCCAGAGCCCTGCGCATGCGCTCGGGTTCTG 123
 OY 124 GAAGACGCGTGATGATGCAAGAGGATTTGCTGCTCTTCTATCTTCTC 183
 DB 124 GAGGACGCGTGATGATGCAAGAGGATTTGCTGCTCTTCTATCTTCTC 183
 OY 184 TTGGTAAAGTCATGATGACCTTACGTCGCGGCTTCGCGACCTGATGGGTACTTCG 243
 DB 184 TTGGTAAAGTCATGATGACCTTACGTCGCGGCTTCGCGACCTGATGGGTACTTCG 243

DB 184 TTGGTAAAGTCATGATGACCTTACGTCGCGGCTTCGCGACCTGATGGGTACTTCG 243
 OY 244 GGGATGTACATGTCACGACGACTGCTCAACTGCACTGATGATGATGAGGACGCGAC 303
 DB 244 GGGATGTACATGTCACGACGACTGCTCAACTGCACTGATGATGATGAGGACGCGAC 303
 OY 304 ATGATCATGACACACCCCGGGTGGCTGCTGCTGCTGCGGAGAACACTCTTCCGCTGC 363
 DB 304 ATGATCATGACACACCCCGGGTGGCTGCTGCTGCTGCGGAGAACACTCTTCCGCTGC 363
 OY 364 TTGGTAAAGTCATGATGACCTTACGTCGCGGCTTCGCGACCTGATGGGTACTTCG 423
 DB 364 TTGGTAAAGTCATGATGACCTTACGTCGCGGCTTCGCGACCTGATGGGTACTTCG 423
 OY 424 CGACGCCACGTCGATGATGTCGTTGGGCGGCTGCTTCTGCTGCTATGATGATGAGG 483
 DB 424 CGACGCCACGTCGATGATGTCGTTGGGCGGCTGCTTCTGCTGCTATGATGATGAGG 483
 OY 484 GATCTCTGCGGATCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 543
 DB 484 GATCTCTGCGGATCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 543
 OY 544 GAGACGCGTCAAGGACTGCAATTCATTCATTCGCGGCGCAATTCATTCGCGGCGCAT 603
 DB 544 GAGACGCGTCAAGGACTGCAATTCATTCATTCGCGGCGCAATTCATTCGCGGCGCAT 603
 OY 604 GCTTGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 628
 DB 604 GCTTGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 628

RESULT 5
 AAT12974
 ID AAT12974 standard; DNA; 2433 BP.

XX AC AAT12974;
 XX AC 25-SEP-1996 (first entry)
 XX DE HCV E1 construct HCC166.
 XX KW HCV; E1, E2; disulphide bond cleavage; envelope protein; vaccine; human;
 KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
 KW ss.
 XX OS Hepatitis C virus.
 XX OS W09604385-A2.
 XX PD 15-FEB-1996.
 XX PF 31-JUL-1995; 95WO-EP03031.
 XX PR 29-JUL-1994; 94EP-0870132.
 XX PA (INNO-) INNOGENETICS NV.
 XX PI Bosman F, Buyse M, De Martynoff G, Maertens G;
 XX DR WPI: 1996-129401/13.
 XX PT Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope
 PT proteins - in presence of disulphide bond cleavage agent, to
 PT produce proteins suitable for direct use in vaccines or diagnostic
 PT assays of HCV
 XX
 XX PS Claim 23; Fig 21; 146pp; English.
 XX SS
 CC AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
 CC and E2 protein coding sequence constructs. These sequences are included
 CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
 CC The recombinant proteins can then be isolated using a method of the
 CC invention. In the method, the envelope proteins are purified by

CC carrying out a disulphide bond cleavage, or a reduction step with a
CC disulphide bond cleavage agent, after lysis of recombinant host cells.
CC The constructs containing the purified HCV envelope proteins can be used
CC for vaccinating humans against HCV, for in vitro detection of HCV
CC antibodies in a sample, and in a serotyping assay for detecting one or
CC more serological types of HCV present in a biological sample. The
CC constructs can also be immobilised on a solid substrate and incorporated
CC into a reversed phase hybridisation assay for determining the presence or
CC the genotype of HCV. The new purification method preserves the
CC conformation of the recombinantly expressed E1, E2 and E1/E2, and
CC eliminates contaminating proteins. Antigens isolated using this method
CC are more reactive with human sera than those isolated by known
CC techniques.
XX
XX

SQ Sequence 2433 BP; 434 A; 745 C; 714 G; 540 T; 0 other;

Query Match 95.7%; Score 605.8; DB 17; Length 2433;
Best Local Similarity 98.1%; Pred. No. 3.2e-158;
Matches 613; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 4 TTGGGTAAAGTATCATGATACCTTACGTCGGCTTCGCGACCTCATGGGGTACATTCCG 63
DB 355 TTGGGTAAAGTATCATGATACCTTACGTCGGCTTCGCGACCTCATGGGGTACATTCCG 414
QY 64 CTCGTCGGCGCCCCCTAGGGGGTGTCCAGAGCCCTGGGCGATGGCGGTCTTG 123
DB 415 CTCGTCGGCGCCCCCTAGGGGGTGTCCAGAGCCCTGGGCGATGGCGGTCTTG 474
QY 124 GAAGCGGGGTGAACATGACACAGGGAATTTGCCGTGTTGCTTCTTACTTCTTC 183
DB 475 GAGGAGCGGGGTGAACATGACACAGGGAATTTGCCGTGTTGCTTCTTACTTCTTC 544
QY 184 TTGGCTTTACTGTCCTGTGACCAATTCAGCTTCGCTTATGAGTGGGCAAGTGC 243
DB 535 TTGGCTTTACTGTCCTGTGACCAATTCAGCTTCGCTTATGAGTGGGCAAGTGC 594
QY 244 GGGATTCACATGTCTCAGAACAGCTCTCCAACTCAAGCATTTGTATGAGGAGCGGAC 303
DB 595 GGGATTCACATGTCTCAGAACAGCTCTCCAACTCAAGCATTTGTATGAGGAGCGGAC 654
QY 304 ATGATCATGCACACCCCGGGTGGCCCTGCTGCTTCGGGAGAACAACTTTCGGCTGC 363
DB 655 ATGATCATGCACACCCCGGGTGGCCCTGCTGCTTCGGGAGAACAACTTTCGGCTGC 714
QY 364 TTGGTAGCCCTCAGCCCAAGCTCGCAGTAGAGAACGCCAGCTCCCACTAGACAATA 423
DB 715 TTGGTAGCCCTCAGCCCAAGCTCGCAGTAGAGAACGCCAGCTCCCACTAGACAATA 774
QY 424 CGACGCCACGTCGATTTGCTGTTGGGGGCGCTGCTTCTGTTCCGCTATGACGTGGG 483
DB 775 CGACGCCACGTCGATTTGCTGTTGGGGGCGCTGCTTCTGTTCCGCTATGACGTGGG 834
QY 484 GATCTGTGGGAGTCTGCTCTGCTGCTCCAGCTGTACCACTTCGCTCCGCTCCGGCAT 543
DB 835 GATCTGTGGGAGTCTGCTCTGCTGCTCCAGCTGTACCACTTCGCTCCGCTCCGGCAT 894
QY 544 GAGACGGTCAGAGACTGCAATTTGCTCAATCTATCCGGCACATAACAGTACCCCTATG 603
DB 895 GAGACGGTCAGAGACTGCAATTTGCTCAATCTATCCGGCACATAACAGTACCCCTATG 954
QY 604 GCTTGGGATATGATGATGAACGTGT 628
DB 955 GCTTGGGATATGATGATGAACGTGT 979

RESULT 6
AAV42305
ID AAV42305 standard; cDNA; 673 BP.
XX
AC AAV42305;
XX
DT 26-OCT-1998 (first entry)
XX

DE HCV core protein cDNA insert of plasmid pA10.

XX HBC; core protein; hepatitis D virus; L-HDAg; virus-like particle;
XX infection; hepatitis B virus surface antigen; HBsAg; immunogen;
KW vaccine; ds.

XX Hepatitis C virus.

XX MO9828004-A1.

XX 02-JUL-1998.

XX 24-DEC-1997; 97WO-AU00884.

XX 24-DEC-1996; 96AU-0004341.

XX (QUEE-) QUEENSLAND DEPT HEALTH SAKZEMSKI VIRUS.

XX Gowans EJ, MacNaughton TB;

XX WPI; 1998-377411/32.

XX P-PSDB; AAM62825.

PT Virus-like particle for, e.g. treating microbial infection -
PT comprises polypeptide from microorganism and sequence from Hepatitis
PT D virus large protein, partially enveloped by Hepatitis B surface
XX antigen

XX Example; Fig 4; 72pp; English.

XX This nucleotide comprises the hepatitis C virus (HCV) cDNA insert
XX in plasmid pA10. This cDNA was obtained by PCR amplification of
XX HCV cDNA using primers (see AAV38849-50) designed to amplify the
XX HCV core gene. The PCR product was cloned into pBluescript KS to
XX create pA10. It encodes a 224-amino acid polypeptide (see AAM62825).
XX The HCV core protein can be used as an immunogen in novel fusion
XX proteins (see AAM62657-59) that comprise HCV core protein and at
XX least 19 amino acids (see AAM62827) of the C-terminal sequence of the
XX large protein from hepatitis D virus (L-HDAg). In novel virus-like
XX particles of the invention, a fusion protein immunogen is at least
XX partially enveloped by hepatitis B surface antigen. The virus-like
XX particle is used to ameliorate or protect against infections caused
XX by hepatitis B virus and/or another microorganism, especially HCV.

SQ Sequence 673 BP; 115 A; 208 C; 188 G; 162 T; 0 other;

Query Match 89.2%; Score 564.8; DB 19; Length 673;
Best Local Similarity 94.1%; Pred. No. 5.1e-147;
Matches 587; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 7 GGTAAAGTATGATACCTTACGTCGGCTTCGCGACCTCATGGGGTACATTCCGCTC 66
DB 1 GGTAAAGTATGATACCTTACGTCGGCTTCGCGACCTCATGGGGTACATTCCGCTC 60
QY 67 GTTCGGGCCCCCTAGGGGGTGTCCAGAGCCCTGCGCATGGCTCCGGGTTCTGAA 126
DB 61 GTTCGGGCCCCCTAGGGGGTGTCCAGAGCCCTGCGCATGGCTCCGGGTTCTGAG 120
QY 127 GACGGGCTGAACATGACACAGGGAATTTGCTGCTTCTTCTATCTTCCCTTG 186
DB 61 GTTCGGGCCCCCTAGGGGGTGTCCAGAGCCCTGCGCATGGCTCCGGGTTCTGAG 120
QY 127 GACGGGCTGAACATGACACAGGGAATTTGCTGCTTCTTCTATCTTCCCTTG 186
DB 121 GACGGGCTGAACATGACACAGGGAATTTGCTGCTTCTTCTATCTTCCCTTG 180
QY 187 GCTTACTGTCCTGTGACCAATTCAGCTTCGCTTATGAGTGGGCAAGCTCCGGG 246
DB 181 GCTTACTGTCCTGTGACCAATTCAGCTTCGCTTATGAGTGGGCAAGCTCCGGG 240
QY 247 ATGTACCATGTCTCAGAACAGCTCTCCAACTCAAGCATTTGTATGAGGAGCGGAC 306
DB 241 GTGTACCATGTCTCAGAACAGCTCTCCAACTCAAGCATTTGTATGAGGAGCGGAC 300
QY 307 ATCATGCACACCCCGGGTGGCCCTGCTGCTTCGGGAGAACAACTTTCGGCTCTG 366
DB 301 ATCATGCACACCCCGGGTGGCCCTGCTGCTTCGGGAGAACAACTTTCGGCTCTG 360

OY	367	GTAAGGCTACCCCCACGCTGCAGTGAAGCCAGCGTCGCCACTAGACAATACGA	426
Db	361	GTACGGCTCACCCCCACGCTGGGGCCAGAAACCTTAGCCTCTACTACGGCAATAACGA	420
OY	427	CGCCACGTCGATTTTGCTGTGGGGGGCGGCTTCTGTTCGGCATATGACGTGGGGGAT	486
Db	421	CGCCACGTCGATTTTGCTGTGGGGGGCGGCTTCTGTTCGGCATATGACGTGGGGGAT	480
OY	487	CTTCGGGATATGTCTTCCTCGTGTCCACTGTTCACATCTTCGCGCTGGCGGGCATGAG	546
Db	481	CTTCGGGATATGTCTTCCTCGTGTCCACTGTTCACATCTTCGCGCTGGCGGGCATGAG	540
OY	547	ACGGTCAGAGACTGCATTTGGCTCAATCTATCCGGCCACATAACAGGTACCCTATGGCT	606
Db	541	ACGGTCAGAGACTGCATTTGGCTCAATCTATCCGGCCACATAACAGGTACCCTATGGCT	600
OY	607	TGGGATATGATGATGAACCTGTAA	630
Db	601	TGGGATATGATGATGAACCTGTCA	624

RESULT 7
ABA03491
ID ABA03491 standard; DNA; 2187 BP.

AC ABA03491;

DT 15-MAR-2002 (first entry)

DE Cuticle protein 1 and 2 secreting hepatitis C virus related DNA #1.

cuticle protein 1; cuticle protein 2; hepatitis C virus; ds.

OS Unidentified.

FH	key	Location/Qualifiers
ET	CDS	1 3197

```

/*tag= a
/product= "AAM4736A"

```

```
FT      /partial
FT      /note="no stop codon"
```

KR97065713-A.

PD 13-OCT-1997.

19-MAR-1996; 96KR-0007404.

PR 19-MAR-1996; 96KR-0007404.

PA (GLDS) LG CHEM LTD.

PI Choo SH, Lee IH, Ryoo WS;

DR WPI: 1998-492654/42.

[illegible]

PT NOAbstract -

PS Disclosure; page 2-4; 7pp; Korean.
xy

CC The present invention relates to cuticle protein 1 and 2 secreting
CC hepatitis C virus. The present sequence is a coding sequence

CC provided in the exemplification of the invention.

sequence 218/ BP; 406 A; 669 C; 631 G; 481 T; 0 other;

Query Match:	89.2%	Score 564.6:	DB 19:	Length 2187:
Best Local Similarity:	93.8%	Pred. No. 8.7e-147:		
Matches 588;	Conservative 0;	Mismatches 39;	Indels 0;	Gaps 0;

OY	4	TTGGGTAAAGGTCATGCAATACCCCTTACGCGGCGCTTCGCCACACCCATAGGGGTCAATTCCG	63
Db	355	TTGGGTAAAGGTCATGCAATACCCCTTACATGCGGCTTCGCCCATCTCATAGGGGTATATTCCG	414
OY	64	CTCGTGGCGGCCCCCTTAGGGGGGTCTGCCAGAGCCCTGGGCGCATGGCGTCCGGGTTCTG	123
Db	415	CTCGTGGCGGCCCCCTTAGGGGGGGCTGCCAGGGCCCTTGGCGACAGGCTGTCGGGGTTCTG	474
OY	124	GAAAGCGGCGGGAATATGCAACAGGGAAATTGGCTGGTGGCTTCTCTATCTTCTC	183
Db	475	GAGGCGGCGGGAATATGCAACAGGGAACTTGGCGGGTGTCTTCTCTATCTTCTC	544
OY	184	TTGGCTTACTGTCCTGTCGTACACATTCCACGCTCCGCTTATGAGTGGCCCAAGTGTCC	243
Db	535	TTGGCTTGTGCTGCTCTGTTTACCATCCACGCTTCGCTTATGAAAGTGGCAAGGTCTC	594
OY	244	GGGATGTACCATGTGCAGACAGCATGCTCCCAACTCAAGATTGTGTATGAGGACGCGGAC	303
Db	595	GGGATGTACCATGTGCAGACAGCATGCTCCCAACTCAAGATTGTGTATGAGGACGCGGAC	654
OY	304	ATGATCATGCAACCCCCGGGTGGTGGCCCGGCTTGGGGAAACAACCTTCCCGCTGC	353
Db	655	ATGATCATGCAATACCCCCGGGTGGTGGCCCTTGGGGGAGCAACAGCTCCCGTGC	714
OY	364	TGGGTAGGCGTCACCCCCCAGCCCTCGCAGCTAGGAACCCAGCGTCCCACTACGACATA	423
Db	715	TGGGTAGGCGTCACCCCCCAGCGCTGGGGCAGGAATCCAGCGTCCCACTACGACATA	774
OY	424	CGAGCCCACTGCATTTGCTGTTGGGGCGGCTGTTCTGTTCGCTATGTAAGTGGG	483
Db	775	CGAGCCCACTGCATTTGCTGTTGGGGCGGCTGTTCTGTTCGCTATGTAAGTGGG	834
OY	484	GATCTCTCGGATGCTGTTCTCTGTCGCCAGCTGTATTCACATCTCGGCTCGCGGCAAT	543
Db	835	GATCTCTCGGATGCTGTTCTCTGTCGCCAGCTGTTCACATCTTCTGCTCGCGGCAAT	894
OY	544	GAGACGGTGCAGAGCATGCAATTGCTCAATCTATCCGCGCACATTAACAGGTCAACCGTATG	603
Db	895	GAGACGGTGCAGAGCATGCAATGCTCAATCTATCCGCGCATTTATCAGGTCAACCGCATG	954
OY	604	GCTTGGGATATGATGATGAATCTGGTAA	630
Db	955	GCTTGGGATATGATGATGAATCTGGTCA	981

RESULT	8
AAQ43889	
ID	AAQ43889 standard; cDNA to mRNA; 2540 BP

AC AAQ43889;

DT 21-OCT-1993 (first entry)
xx

NANB hepatitis virus polynucleotide N-2540-2.

KM Non-A, non-B; virus; polymerase chain reaction; detection;
KM sensitive; specific; HCV; NANBH; CC

XX Non-A, non-B hepatitis virus.

XX	Key	Location/Qualifiers
FH		

CDS	342..2340
ET	/*taq= a
ET	

FTT	1.1341	/*tag=	b
FTT			
FTT			

XX 13

XX :
 :
 :
 :
 :
 :
 :
 :

XX

XX

PR 12-JUN-1990; 90JP-0153401.
PR 08-NOV-1990; 90JP-0304405.
PA (NAKA/) NAKAMURA T.
XX
XX
XX WPI: 1993-199637/25.
DR P-PSDB; AAR38279.
XX
XX Antigen related to non-A and non-B hepatitis virus - comprises
PT non-translation region comprising 340 - 341 mols. of nucleotides,
PT non-translation region comprising 1885 - 2551 mols. of
PT nucleotides including region 1,149 and, etc.
XX
XX
XX Claim 3; Page 19-20; 73pp; Japanese.
XX
XX The sequence is that of NANB hepatitis virus polynucleotide N-2540-2
CC which codes for a non-A, non-B (NANB) hepatitis virus gene HC-OM.
CC The polypeptide it encodes may be used in a system for detecting
CC NANB hepatitis. This method is highly specific and sensitive, and
CC can detect NANB hepatitis virus which could not be detected by
CC conventional methods.
XX
XX
SQ Sequence 2540 BP; 471 A; 775 C; 741 G; 553 T; 0 other;

Query Match 89.2%; Score 564.6; DB 14; Length 2540;
Best Local Similarity 93.8%; Pred. No. 9,1e-147;
Matches 588; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
OY 4 TTGGGTAAGTCATGATGATCCCTTACGTCGCGCTTCCGACCTCATGGGGTATCCG 63
DB 696 TTGGGTAAGTCATGATGATCCCTTACATGCGGCTTCCGATCTCATYGGGATATTCG 755
OY 64 CTCGTGCGGCCCCCTAGGGGGTGTGCTGCCAGAGCCCTTGGCATGGCGGTTCG 123
DB 756 CTCGTGCGGCCCCCTAGGGGGTGTGCTGCCAGAGCCCTTGGCATGGCGGTTCG 815
OY 124 GAAGAGCGGCTGATATGATCAACAGGAAATTTGCTGCTTCTTCTATCTTCTC 183
DB 816 GAGGAGCGGCTGATATGATCAACAGGAAATTTGCTGCTTCTTCTATCTTCTC 875
OY 184 TTGGCTTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 243
DB 876 TTGGCTTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 935
OY 244 GGGATGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 303
DB 936 GGGATGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 995
OY 304 ATGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 363
DB 996 ATGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1055
OY 364 TTGGGTAAGTCATGATGATCCCTTACGTCGCGCTTCCGACCTCATGGGGTATCCG 423
DB 1056 TTGGGTAAGTCATGATGATCCCTTACATGCGGCTTCCGATCTCATYGGGATATTCG 1115
OY 424 CGAGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 483
DB 1116 CGAGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1175
OY 484 GATCTCTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 543
DB 1176 GATCTCTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1235
OY 544 GAGAGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 603
DB 1236 GAGAGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1295
OY 604 GCTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 630
DB 1296 GCTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1322

RESULT 9
AA063753
ID AA063753 standard; cDNA to mRNA; 2540 BP.
XX
XX AA063753;
AC
XX
XX 30-JAN-1995 (first entry)
DT
XX
XX NANBH genomic fragment #2.
DE
XX
XX Polymerase chain reaction; PCR; primer; amplify; detection; NANBH;
KM non-A, non-B hepatitis virus; 5'-terminal region; core protein; ss.
XX
XX
XX Synthetic.
OS
XX
XX JP06125777-A.
PN
XX
XX 10-MAY-1994.
PD
XX
XX 20-JUN-1991; 91JP-0247120.
PF
XX
XX 20-JUN-1991; 91JP-0247120.
PR
XX
XX (NAKA/) NAKAMURA T.
PA
XX
XX WPI: 1994-187937/23.
DR
XX
XX Oligonucleotide primer pairs specific for non-A, non-B hepatitis
PT virus - used for high sensitivity detection of non-A non-B (NANB)
PT hepatitis virus
PS
XX
XX Disclosure; Page 24-25; 25pp; Japanese.
XX
XX The sequences given in AA063753-53 represent fragments of the non-A,
CC non-B hepatitis virus (NANBH) genome. These fragments were amplified
CC using the primers given in AA063732-51. These primers were used in the
CC detection of NANBH. The primers are based on the 5'-terminal region and
CC the core protein coding region. The method allows highly sensitive
CC detection of NANBH.
XX
XX
SQ Sequence 2540 BP; 470 A; 775 C; 742 G; 553 T; 0 other;

Query Match 89.2%; Score 564.6; DB 15; Length 2540;
Best Local Similarity 93.8%; Pred. No. 9,1e-147;
Matches 588; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
OY 4 TTGGGTAAGTCATGATGATCCCTTACGTCGCGCTTCCGACCTCATGGGGTATCCG 63
DB 696 TTGGGTAAGTCATGATGATCCCTTACATGCGGCTTCCGATCTCATYGGGATATTCG 755
OY 64 CTCGTGCGGCCCCCTAGGGGGTGTGCTGCCAGAGCCCTTGGCATGGCGGTTCG 123
DB 756 CTCGTGCGGCCCCCTAGGGGGTGTGCTGCCAGAGCCCTTGGCATGGCGGTTCG 815
OY 124 GAAGAGCGGCTGATATGATCAACAGGAAATTTGCTGCTTCTTCTATCTTCTC 183
DB 816 GAGGAGCGGCTGATATGATCAACAGGAAATTTGCTGCTTCTTCTATCTTCTC 875
OY 184 TTGGCTTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 243
DB 876 TTGGCTTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 935
OY 244 GGGATGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 303
DB 936 GGGATGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 995
OY 304 ATGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 363
DB 996 ATGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1055
OY 364 TTGGGTAAGTCATGATGATCCCTTACGTCGCGCTTCCGACCTCATGGGGTATCCG 423
DB 1056 TTGGGTAAGTCATGATGATCCCTTACATGCGGCTTCCGATCTCATYGGGATATTCG 1115

This invention describes a novel hepatitis C virus (HCV) cell culture system comprising human hepatoma cells that contain an integrated HCV-RNA construct (I). (I) contains the HCV-specific RNA segments 5'-NTR (non-translated region), NS (non-structural)3, NS4A, NS4B, NS5A, NS5B and 3'-NTR, and a selectable (marker) gene (II). The cell cultures, and/or (I), are used to prepare, evaluate and/or test therapeutic and/or diagnostic agents for HCV infections, and to prepare vaccines against HCV infection (particularly preparation of attenuated HCV). The can also be used for preparation of a liver-specific delivery system for gene therapy, and to identify cells permissive for HCV replication. Virus RNA replicates autonomously and with high efficiency in this cellular system, so that variations in replication rates can be measured (for screening antiviral agents) quantitatively or qualitatively, using standard laboratory equipment. Efficient replication of HCV RNA is only achieved when the specified RNA segments are present and when the transfected cells are maintained under permanent selection pressure.

SQ Sequence 11076 BP; 2221 A; 3297 C; 3149 G; 2409 T; 0 other;

Query Match	89.2%	Score 564.6	DB 21	Length 11076
Best Local Similarly	93.8%	Pred. 1.5e-146		
Matches 588	Conservative	0	Mismatches 39	Indels 0
				Gaps 0

QY	4	TTTGSTGTAAGTCATTCGATTAACCTTTAGCTGCGGCTTGCGCCAGACTATGGGGTAATTTCCG	63
Db	2167	TTTGGGTAAAGTCATTCGATTAACCTTCACGTGCGGGCTTGCGCGCATCTATGGGGTAATTTCCG	2226
QY	64	CTCTCGCGGCCCCCCCCCTAGAGGGGGTGTCTGCACAGCCCTTGCCCATAGCGGTCCGGGGTTCTG	123
Db	2227	CTGTCTGGGCGCCCCCTTAGGGGGGGCTGTGCACAGGGCCCTTGCGCATAGCGGTCCGGGGTTCTG	2286
QY	124	GAAGACGGCGTGAACATATGCAACAGGGAATTGGCTGTGCTCTTTCTATCTTCTCCTC	183
Db	2287	GAGAGACGGCGGTAACATATGCAACAGGGAATTCGCCGGTGTCTCTCTTCTATCTTCTCTT	2346
QY	184	TTTGCTTTTACGTCGCTGTGTGACCAATTCACAGCTCCGCGTTATGAGGTGGCAAGCTGTC	243
Db	2347	TTTGCTTTTGTCTGTCTCTTTTGACCAATTCACAGCTCCGCTTATGAGAGTGGCAAGCTATTC	2406
QY	244	GGGATGTACCATGTCAAGAACAGCACTGCTCCAACTCAAGCATTTGTATAGGCAACCGGAC	303
Db	2407	GGATGTATACCATGTCAAGAACAGCACTGCTCCAAAGCAATTTGTATAGGCAACCGGAC	2466
QY	304	ATGATCATGCACACCCCCGGGGTGGTCCCTGTGGTTCCGGGAACAACACTCTTCCCGCTGC	363
Db	2467	ATGATCATGTGATTAACCCCCGGGGTGGTCCCTGTGGTTCCGGGAACAACACTCTTCCCGCTGC	2526
QY	364	TGGGTACAGGCTCAACCCCCACAGCTCGGAGCTAGGAAGCGCAGGGTCCCACTACGACATA	423
Db	2527	TGGGTACAGGCTCACTCCCAAGCTCGGGGCGAGGAACGCTAGGGTCTCCCACTACGACATA	2566
QY	424	CGACGCCACGTCGATTTGTCGTGTGGGGGGCGGCTCTTCTGTTCGCGTATGATACGTGGGG	483
Db	2587	CGACGCCACGTCGATTTGTCGTGTGGGGGGCGGCTCTTCTGTTCGCGTATGATACGTGGGA	2646
QY	484	GATCTCTGGGAGATCTGCTCTCTCTCTCTCCAGCTTTCACCATCTCGGCTTGCCGGGAT	543
Db	2647	GATCTCTGGGAGATCTGTTTCTCTCTCTCTCCAGCTTTCACCATCTCTCGCTTGCCGGGAC	2706
QY	544	GAGACGGTGGAGAGCTCAATTTGCTCAATCTATCCGGGCAATANAAGTACCCGATG	603
Db	2707	GAGACATTAACAGAGCTCAATTTGCTCAATATATCCGGGCAAGTGAACAGTACCGGATG	2766
QY	604	GCTTGGGATATGATGATGAACCTGTAA	630
Db	2767	GCTTGGGATATGATGATGAACCTGTAA	2793

RESULT	13
AAQ24467	
ID	AAQ24467 standard; DNA; 1880 BP.
XX	
AC	AAQ24467;

XX	09-NOV-1992	(first entry)
DT		
XX		
DE		
XX		
KW		
polymerase chain reaction; vaccine; antibody; ss.		
non-A, non-B hepatitis virus; NANBHV; PCR; amplification		
NANB hepatitis virus strain HC-J4 genome.		

	Location/Qualifiers
FH	342..1880
FT	/*tag= a
FT	/label= HC-J4
FT	

PN	EP485209-A.	
XX		
PD	13-MAY-1992.	
XX		
PF	07-NOV-1991;	91EP-0310297.
XX		
PR	08-NOV-1990;	90JP-0304405.

PT Recombinant CDNA of NANBH virus strain HC-J5 and corresp.
PT peptides - useful for diagnosis and in vaccines and immunological
PT pharmaceuticals

CC This sequence is the genome of the non- η , non-B hepatitis virus
CC (MANBHV) strain HC-74. This sequence was derived by amplification
CC by polymerase chain reaction. The nucleotide sequences derived from
CC this amplification can be used to detect MANBHV infection which could
CC not be detected by conventional methods. The detection kits allow
CC highly specific and sensitive detection at an early phase of
CC infection. The polypeptide product of this coding sequence can be used
CC for the manufacture of vaccines and immunological pharmaceuticals
CC and also to produce antibodies specific to MANBHV.

Sequence 1880 BP; 333 A; 593 C; 556 G; 398 T; 0 other;

Query Match	88.9%	Score 563;	DB 13;	Length 1880;
Best Local Similarity	93.6%	Pred. No. 2.3e-146;		
Matches 587; Conservative	0;	Mismatches 140;	Indels 0;	Gaps 0;

QY	4	TTTGGTAAGGTCATGCAATACCTTTACGGGGGGCTTCCGACCACTCATGGGGTAATATCCG	63
Db	636	TTGGTTAAGGTCATGCAATACCTTTACGGGGGGCTTCCGACCTCATAGGGGTAATATCCG	755
QY	64	CTCGTCGGCGCCCCCTTAAGGGGGTGTCCAGAGCCCTTGCGGCATGCGCGTGGGGTCTCG	123
Db	756	CTCGTCGGCGCCCCCTTAAGGGGGGGTCCAGAGCCCTTTGGCACACGGTGTCCGGGTTCTG	815
QY	124	GAAGCGCGCGTAATATGCAACAGGGAAATTGGCTGGTGTCTCTTCTCTATCTTCTCTC	183
Db	816	GAGGCGCGCGTAATATGCAACAGGGAACTTGGCCGGGTGCTCTTCTCTATCTTCTCTC	875
QY	184	TTTGGTTTACTGTCCTGTCTACACATTCACAGCTTCCGCTTATGAGGCGCCACAGTGTCC	243
Db	876	TTGGCTTTTGCTGTCCTGTTTBACCATCCACAGCTTCCGCTTATGAGGCGCCACAGTGTCC	935
QY	244	GGGATGTACCATGTTCACGAACGACTGCTCCCAATCAAGACATTTGTATGAGGACGGGAC	303
Db	936	GGGATATTCACATGTTCACGAACGAGTGTCCCAATCAAGACATTTGTATGAGGACGGGAC	995
QY	304	ATGATCATGCACACCCCGGGGTGGGTGGCCGTCTGGGGGGAACAACCTTCCCGGGTGC	363

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 18, 2003, 05:52:15 ; Search time 1691 Seconds
(without alignments)
6062.533 Million cell updates/sec

Title: us-09-899-303a-7

Perfect score: 633
Sequence: 1 ATGTTGGGTAAGTCATCGA.....TGATGATGAACGTGTAATAG 633Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estb:*
2: em_esthum:*
3: em_estln:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_estl:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vtc:*
22: em_gss_fun:*
23: em_gss_man:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
C 1	75.4	11.9	488 10	AV755731 AV755731
C 2	61	9.6	492 10	AV758366 AV758366
C 3	39.8	6.3	1270 13	BG968359 BG968359
C 4	39.6	6.3	423 9	AA459034 AA26b11.r
C 5	39	6.2	359 13	BJ252669 BJ252669
C 6	39	6.2	375 13	BJ246716 BJ246716

7	38.6	6.1	399	10	AV638521	AV638521
8	38.6	6.1	434	10	AV637507	AV637507
9	38.6	6.1	440	10	AV637983	AV637983
10	38.6	6.1	450	10	AV637259	AV637259
11	38.6	6.1	451	10	AV637328	AV637328
12	38.6	6.1	451	10	AV637643	AV637643
13	38.6	6.1	453	10	AV634724	AV634724
14	38.6	6.1	454	10	AV637050	AV637050
15	38.6	6.1	456	10	AV635382	AV635382
16	38.6	6.1	473	10	AV632765	AV632765
17	38.6	6.1	481	10	AV635503	AV635503
18	38.6	6.1	485	10	AV632811	AV632811
19	38.6	6.1	506	10	AV392445	AV392445
20	38.6	6.1	508	10	AV634095	AV634095
21	38.6	6.1	526	10	AV641895	AV641895
22	38.6	6.1	533	10	AV638125	AV638125
23	38.6	6.1	537	10	AV632335	AV632335
24	38.6	6.1	588	10	AV387329	AV387329
25	38.6	6.1	525	10	BE337089	BE337089
26	38.4	6.1	671	13	BI723733	BI723733
27	38.2	6.0	431	10	AV639153	AV639153
28	38.2	6.0	501	10	AV638474	AV638474
29	38	6.0	1195	17	BI09802	BI09802
30	37.8	6.0	435	14	C72860	C72860
31	37.6	5.9	222	13	BM256206	BM256206
32	37.6	5.9	456	13	BI776179	BI776179
33	37.6	5.9	478	12	BF775654	BF775654
34	37.6	5.9	486	13	BI776135	BI776135
35	37.6	5.9	570	13	BI776136	BI776136
36	37.6	5.9	593	13	BI849151	BI849151
37	37.4	5.9	497	10	AV633658	AV633658
38	37.4	5.9	624	13	BI723734	BI723734
39	37.2	5.9	925	17	CNS00919	CNS00919
40	37	5.8	945	17	CNS05KMC	CNS05KMC
41	37	5.8	970	17	CNS010C9	CNS010C9
42	36.8	5.8	431	10	AV636681	AV636681
43	36.8	5.8	439	10	AV638045	AV638045
44	36.8	5.8	477	10	AV635762	AV635762
45	36.8	5.8	504	10	AV637539	AV637539

ALIGNMENTS

RESULT 1
LOCUS AV755731 488 bp mRNA linear EST 19-OCT-2000
DEFINITION AV755731 BM Homo sapiens CDNA clone BMFAK803 5', mRNA sequence.
ACCESSION AV755731
VERSION AV755731.1 GI:10913579
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 488)
Gu,J., Zhao,M., Huang,Q., Xu,X., Li,Y., Peng,Y., Song,H., Xiao,H.,
L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,H., Lu,G.,
Yang,Y., Gao,G., Wang,Z., Zhang,Q., Chen,S., Han,Z. and Chen,Z.
Homo sapiens CDNA BM clones
Unpublished (2000)
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
location/Qualifiers
1..488
/organism="Homo sapiens"

FEATURES

source

```

/db_xref="taxon:9606"
/clone="BMFAK03"
/clone_lib="BM"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"
/lab_host="BM25.8"
/note="Vector: pTriplex2; Site_1: sflIA; Site_2: sflIB"

BASE COUNT      116 a      134 c      137 g      97 t      4 others

ORIGIN
Query Match      11.9%; Score 75.4; DB 10; Length 488;
Best Local Similarity 70.3%; Pred. No. 1.7e-09;
Matches 130; Conservative 0; Mismatches 51; Indels 4; Gaps 2;

OY  445 GTTGGGGGGGGCGCTTTCTGTCTCCCTATATACGTGGGGGAGCTCTCGGGAGTCTGTTC 504
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB  472 GTGGGTGACACTCCCTCTGCTGACCTCTCTAAGTGTGGACCTTCGACAGCAAGTATG 413
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY  505 CTGCTCTCCACCTTTTACACATCTTCGCTCGCGTCCGGCATGAGACGGTGCAGACTGCAAT 564
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB  412 CTTCGAGTTTCACTG---ATCATCTGCGCTTCAGCACACATGAGTTTGTGATATATGCAAC 356
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY  555 TGTCTCAATCTATCCCGGCGCACATAACAGGTACCGCTATG-TCGTTGGATATGATGATGAA 623
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB  355 TGTCTCAATCTATCTGCGGCCCATCTACTGTGACACCGTATGATGATGATGATGATGAA 296
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY  624 CTGGT 628
      ||| | |
DB  295 CTGGT 291

RESULT 2
AV758366/c
LOCUS      AV758366      492 bp      mRNA      linear      EST 19-OCT-2000
DEFINITION AV758366 BM Homo sapiens CDNA clone BMFAK03 5', mRNA sequence.
ACCESSION  AV758366
VERSION    AV758366.1 GI:10916214
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 492)
            Gu,J., Zhao,N., Huang,O., Xu,X., Li,Y., Peng,Y., Song,H., Xiao,H.,
            Gu,Y., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,
            ,L., Xu,S., Gu,W., Tu,X., Jia,J., Fu,G., Ren,S., Zhong,W., Lu,G.,
            Yang,Y., Gao,G., Wang,Z., Zhang,Q., Chen,S., Han,Z. and Chen,Z.
            Homo sapiens CDNA BM clones
            Unpublished (2000)
            Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919 (ex.45)
            Fax: 86-21-50801922
            Email: hanzgehc.sh.cn
            This clone is available at CHGC in Shanghai.

FEATURES
            source
                1..492
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="BMFAK03"
                /clone_lib="BM"
                /tissue_type="Bone marrow"
                /cell_type="CD34+ hematopoietic stem/progenitor cell"
                /lab_host="BM25.8"
                /note="Vector: pTriplex2; Site_1: sflIA; Site_2: sflIB"

BASE COUNT      124 a      128 c      125 g      112 t      3 others

ORIGIN
Query Match      9.6%; Score 61; DB 10; Length 492;
Best Local Similarity 68.0%; Pred. No. 1.1e-05;
Matches 115; Conservative 0; Mismatches 50; Indels 4; Gaps 2;

```

OY	461	TCTGTCGGCATATGCAGCGGGGAAATCTCTCCGGGAATCTCCTCGTCCCAACTGT	520
Dd	457	TCTGATACAGCTCACACAGTGTTGAGACCTCTCGCTTGGGGTAATCGCTTGACGCCAACATG-	399
OY	521	TCACCATCTCGCCTCGCCGCGCATGAGACGGTGCAGAGCTGACAATTGCTCAATCTATCCCG	580
Dd	398	- - ATATATCTCTCAGACAGCAACATTTGGTTGTGTCAAGAATGCAACATGCTATCTATCCGG	341
OY	581	GCCACATAACAGATGC-ACCAGTATGGCTTGGGATATGATGATGATCAACTGGT	628
Dd	340	GCTGCATCATCTGACTACAGTATGAGCATAGGCTATGATGATGATCAACTGGT	292
 RESULT 3 BG968359/c			
LOCUS	DEFINITION	602834809P1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:498195 5'	
ACCESSION	VERSION	BG968359	
KEYWORDS	SOURCE	BG968359.1 GI:14355996 EST.	
ORGANISM	house mouse. Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
TITLE	JOURNAL	NIH-MGC http://mgc.nci.nih.gov/ National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)	
COMMENT	CONTACT: Robert Strausberg, Ph.D. Email: cgabbs-remail.nih.gov Tissue Procurement: Jeffrey E. Green, M.D. cNSA Library Preparation: Life Technologies, Inc. cNSA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.llnl.gov Plate: L1AM11003 row: a column: 04 High quality sequence start: 6 High quality sequence stop: 416.		
 FEATURES			
source	location/Qualifiers	1..1270	
	/organism="Mus musculus"		
	/strain="FVB/N"		
	/db_xref="taxon:10090"		
	/clone="IMAGE:498195"		
	/clone_1id="NCI_CGAP_Co24"		
	/lab_host="DH10B (TI phage-resistant)"		
	/note="Organ: colon; Vector: pCMV-SPORT6; Site.1: NotI; Site.2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.6 kb. Constructed by life technologies. Note: this is a NCI_CGAP library."		
BASE COUNT	274 a 387 c 445 g 164 t		
ORIGIN			
Query Match	6.3%; Score 39.8; DB 13; Length 1270;		
Best Local Similarity	45.0%; Pred. No. 6.2;		
Matches 149; Conservative	0; Mismatches 182; Indels 0; Gaps 0;		
OY	81	AGGGGTGCTGCCAGAGCCCTGGCGCATGGCTCGGGGTTGGAAGAGCGGCTGACACTA	140
Dd	1125	ACTGGGCCCTTGGGTGCTCGCTGCTCTCTTCGCGCGCTGACATCCCTGCGCCCTC	1066
OY	141	TGCAACAGGGAATTTGGCCCTGGTGCCTTTCTATCTATCTCACTCTTGAGCTTACTGTCG	200
Dd	1065	TGCTGACATCTGTCTCTCTGCTGACATACCTCCCGCTGTCCCGGAGTGTCTGTGCCCCG	1006
OY	201	TCTGACCATTTCCAGCTTCGCTTAGAGGTGGGCAACGATCCGGGATGTACACTGCAC	260
Dd	1005	TCCCATCTGTCGCGGATCCGCTGTGGGGGTCCCAATCCTCGGGGTGCTTCTCTGTCTAG	946

[illegible]

RESULT 4					
AA459034	AA459034	423 bp	mRNA	linear	EST 13-AUG-1997
LOCUS					
DEFINITION	aa26b11.r1 NC1_GCAP_GCB1 Homo sapiens cDNA clone IMAGE:814365 5',				
	mRNA sequence.				
ACCESSION	AA459034				

BASE COUNT ORIGIN	66 a	162 c	101 g	94 t
----------------------	------	-------	-------	------

0y 466 GCGATTATACGGGGGGAATCTGTGGGAATCTGTCTCCCTGTCCACAGCTGTACACAT 527
Db 203 CATCAGGGGAGCAGGTGACTTCAAGGAACTCCCGGAGACCTGGCTAGACGTTAAACAG 265
0y 528 CTCGGCTGGCCGGCATGAGAGGTGACAGACTGCATATGCTCATATATCCCGGAC 585
Db 265 CACGTCCTTCCACAGCTGAGCCAGCCAGCTTTGGGAGAGGGCTTCGTGCCCCCCCAC 320

RESULT 5	
BJ252669/c	
LOCUS	359 bp mRNA linear
DEFINITION	BJ252669 Y. Ogihara unpublished cDNA library, Wh_f Triticum aestivum
ACCESSION	BJ252669
VERSION	BJ252669.1
KEYWORDS	GI:20061830
SOURCE	EST.
	bread wheat.

RESULT
PJ246716

LOCUS	Bj246716	375 bp	mRNA	linear	EST 05-APR-2000
DEFINITION	Bj246716 y. Ogihara unpublished cDNA library, wh_f Triticum aestivum cDNA clone whf25g19 5', mRNA sequence.				
ACCESSION	Bj246716				
VERSION	Bj246716.1	GI:20058228			
KEYWORDS	EST.				
SOURCE	bread wheat.				
ORGANISM	Triticum aestivum				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticaceae; Triticum.				
AUTHORS	1 (bases 1 to 375)				
JOURNAL	Ogihara,Y. and Murai,K. Expressed genes in Triticum aestivum Unpublished (2002) Contact: Tadasu Shin-I Center For Genetic Resource Information National Institute of Genetics 1111 Yata, Mishima, Shizuoka 411-8540, Japan Tel: 81-559-81-6856 Fax: 81-559-81-6855 Email: tshini@genes.nig.ac.jp.				
FEATURES	location/Qualifiers				
source	1..375				
	/organism="Triticum aestivum"				
	/cultivar="Chinese Spring"				
	/db_xref="taxon:4565"				
	/clone="whf25g19"				
	/clone_1ib="Y. Ogihara unpublished cDNA library, wh_f"				
	/tissue.type="spike at flowering date"				
	/dev_stage="Peekes' scale 10.5.1"				
	/note="Vector: lambda Uni-ZAP XR, excised phagmid; Site_1: EcoRI; Site_2: XhoI; Plants were grown under hydroponic conditions at UC Davis, salt stressed for 12 hours, and for 7 days, then dissected and frozen (Akhunov in J Dvorak Lab). Total RNA was prepared from sheath tissue, equal quantities of RNA were pooled from the two samples, polyA was purified from the pooled RNA, a cDNA library was made, and the cDNA clones were in vivo excised to give pluescript phagemids in the TJ Clouse lab at the University of California, Riverside (Akhunov, Chin , Choi, Clouse, Penton, Kianian, Otto, Simons, Zhang). Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."				
BASE COUNT	81 a 107 c 110 g 77 t				
ORIGIN					
Query Match	6.2%; Score 39; DB 13; Length 375;				
Best Local Similarity	58.0%; Pred. No. 7.1;				
Matches	69; Conservative 0; Mismatches 50; Indels 0; Gaps 0;				
Qy	24 CCTTACGTGCGGGCTTCGCCAGCCTCATGGGGGTACTTCGGCTGCGGCGCCCTTAGG 83				
Db	36 CTTCGAATGCACAACAGCGCGCTGTGAAGAAGCGGCTCAAGGCGGTCGACAGCGCGTGCGG 95				
Qy	84 GGCTCTGCCAGACGCCCTGCGCGCATGCGGTTCGCAAGACGGCGTAACATAG 142				
Db	96 GGACGCGGCGACGCCCTGTGGCGCCAGAGAGCTGACGCTGCGCGGTGCACAGTCCCAAAG 154				
RESULT 7					
LOCUS	AV638521	399 bp	mRNA	linear	EST 15-DEC-2000
DEFINITION	AV638521 Chlamydomonas reinhardtii 5% CO2 chlamydomonas reinhardtii				
ACCESSION	AV638521				
KEYWORDS	AV638521.1 GI:10781841				
SOURCE	EST.				
ORGANISM	Chlamydomonas reinhardtii. Chlamydomonas reinhardtii Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales; Chlamydomonadales; Chlamydomonas.				
REFERENCE	1 (bases 1 to 399)				

```

AUTHORS      Asamizu,E., Miura,K., Kuchio,K., Inoue,Y., Fukuzawa,H., Ohyama,K.,
              Nakamura,Y. and Tabata,S.
TITLE        Generation of expressed sequence tags from low-CO2 and high-CO2
              adapted cells of Chlamydomonas reinhardtii
JOURNAL      DNA Res. 7 (5), 305-307 (2000)
MEDLINE      20539644
COMMENT      Contact: Erika Asamizu
              The First Laboratory for Plant Gene Research
              Kazusa DNA Research Institute
              Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
              Email: asamizue@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES     source
              1..399
              /organism="Chlamydomonas reinhardtii"
              /strain="Cg"
              /db_xref="taxon:3055"
              /clone="HC087d07_-r"
              /note="Vector: pluescriptII SK-, Site_1: EcoRI; Site_2:
              XhoI; The cDNA library was constructed from cells cultured
              in a medium with bubbling air containing 5% carbon
              dioxide"
BASE COUNT   70 a          144 c          122 g          63 t
ORIGIN
Query Match    Best Local Similarity 6.1%; Score 38.6; DB 10; Length 399;
Matches 92; Conservative 0; Mismatches 89; Indels 0; Gaps 0;
QY 183 CTTGGCTTTACGTCGCCGTCTGCAGCATTCGCAGCTCCGCGTATGAGTGCGCAAGCGTCTC 242
Db 44 CTTCGACCCCCCTGACGCGCTGCTCCATCGTCGACACCACAATCTCGCGTGGAACCATCTT 103
QY 243 CGGAGTAGTACCATGTCACAGAAGACTGCTCCAACTCAAGCATTTGTATTGAGCACCGGA 302
Db 104 CGGCGTGTGGCCCGGCGAGCAAGCTGACCAATAATCAACGCGCGCGAGCAAGTGTGCGGG 163
QY 303 CATGATCATGCACACACCCCGGGTGGTCCCTTGCCTTGGGAGAAACAATCTTCCCCTG 362
Db 164 CATGGGCAATCAACGGTCCCGCGACGCGTGTTCGATTCGCCCTGGAAGAGACGCCCGCGTG 223
QY 363 C 363
Db 224 C 224
RESULT 8
AV637507 LOCUS 434 bp mRNA linear EST 15-DEC-2000
DEFINITION AV637507 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
ACCESSION   cDNA clone HC07a01_r 5', mRNA sequence.
VERSION     AV637507
KEYWORDS    AV637507.1 GI:10780827
SOURCE      ESP.
ORGANISM   Chlamydomonas reinhardtii.
REFERENCE   Chlamydomonas reinhardtii.
AUTHORS     Chlamydomonas reinhardtii.
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
            Chlamydomonadaceae; Chlamydomonas.
            1 (bases 1 to 434)
            Asamizu,E., Miura,K., Kuchio,K., Inoue,Y., Fukuzawa,H., Ohyama,K.,
            Nakamura,Y. and Tabata,S.
            Generation of expressed sequence tags from low-CO2 and high-CO2
            adapted cells of Chlamydomonas reinhardtii
            DNA Res. 7 (5), 305-307 (2000)
COMMENT     Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizue@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES     Location/Qualifiers
              1..434
              /organism="Chlamydomonas reinhardtii"

```



```

/strain="C9"
/db_xref="taxon:3055"
/clone="HC074a01_r"
/clone_1lb="Chlamydomonas reinhardtii 5% CO2"
/notes="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from cells cultured
in a medium with bubbling air containing 5% carbon
dioxide"
BASE COUNT      80 a      149 c      130 g      75 t
ORIGIN
Query Match      6.1%; Score 38.6; DB 10; Length 434;
Best Local Similarity 50.8%; Pred. No. 9.5;
Matches 92; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 183 CTTGGCTTTACTGTCGTCTGACCAATTCACGTTCCGCTTATGAGTGCGCAACGTGTC 242
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 21 CTTGACACCCCTGGACGGCTGTCATCGTCGACACCAACTTCGCGTGGGACACATCTT 80

QY 243 CGGATGTACCATGTACAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGCGA 302
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 81 CGGCGTGTGGCCCGCGACAGCTGACCAACATCACCAGCGCGACGAGTGCTGCCGG 140

QY 303 CATGATCATGCACACCCCGGCTGCTCCCTGCTGGGAGAACAACTCTTCCGCTG 362
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 141 CATGGCATCTACGGTCCCGCACCGCTGTCTGATTGCTTGAAGGACGCCCGGCTG 200

QY 363 C 363
Db 201 C 201

RESULT 9
LOCUS      AV637983      440 bp      mRNA      linear      EST 15-DEC-2000
DEFINITION AV637983 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
ACCESSION  AV637983
VERSION     AV637983.1 GI:10781303
KEYWORDS    EST.
SOURCE      Chlamydomonas reinhardtii.
ORGANISM    Chlamydomonas reinhardtii.
REFERENCE   1 (bases 1 to 440)
AUTHORS     Asamizu, E., Miura, K., Kucho, K., Inoue, Y., Fukuzawa, H., Ohyama, K.,
            Nakamura, Y. and Tabata, S.
TITLE       Generation of expressed sequence tags from low-CO2 and high-CO2
            adapted cells of Chlamydomonas reinhardtii
JOURNAL     DNA Res. 7 (5), 305-307 (2000)
MEDLINE     20539644
COMMENT     Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizuekazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
            Location/Qualifiers
                1..440
                /organism="Chlamydomonas reinhardtii"
                /strain="C9"
                /db_xref="taxon:3055"
                /clone_1lb="Chlamydomonas reinhardtii 5% CO2"
                /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
                XhoI; The cDNA library was constructed from cells cultured
                in a medium with bubbling air containing 5% carbon
                dioxide"
BASE COUNT      76 a      156 c      137 g      71 t
ORIGIN
Query Match      6.1%; Score 38.6; DB 10; Length 440;
Best Local Similarity 50.8%; Pred. No. 9.5;
Matches 92; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

```

```

QY 183 CTTGGCTTTACTGTCGTCTGACCAATTCACGTTCCGCTTATGAGTGCGCAACGTGTC 242
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 80 CTTGACACCCCTGGACGGCTGTCATCGTCGACACCAACTTCGCGTGGGACACATCTT 139

QY 243 CGGATGTACCATGTACAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGCGA 302
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 140 CGGCGTGTGGCCCGCGACAGCTGACCAACATCACCAGCGCGACGAGTGCTGCCGG 199

QY 303 CATGATCATGCACACCCCGGCTGCTCCCTGCTGGGAGAACAACTTCCGCTG 362
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 200 CATGGCATCTACGGTCCCGCACCGCTGTCTGATTGCTTGAAGGACGCCCGGCTG 259

QY 363 C 363
Db 260 C 260

RESULT 10
LOCUS      AV637259      450 bp      mRNA      linear      EST 15-DEC-2000
DEFINITION AV637259 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
ACCESSION  AV637259
VERSION     AV637259.1 GI:10780579
KEYWORDS    EST.
SOURCE      Chlamydomonas reinhardtii.
ORGANISM    Chlamydomonas reinhardtii.
REFERENCE   1 (bases 1 to 450)
AUTHORS     Asamizu, E., Miura, K., Kucho, K., Inoue, Y., Fukuzawa, H., Ohyama, K.,
            Nakamura, Y. and Tabata, S.
TITLE       Generation of expressed sequence tags from low-CO2 and high-CO2
            adapted cells of Chlamydomonas reinhardtii
JOURNAL     DNA Res. 7 (5), 305-307 (2000)
MEDLINE     20539644
COMMENT     Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizuekazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
            Location/Qualifiers
                1..450
                /organism="Chlamydomonas reinhardtii"
                /strain="C9"
                /db_xref="taxon:3055"
                /clone_1lb="Chlamydomonas reinhardtii 5% CO2"
                /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
                XhoI; The cDNA library was constructed from cells cultured
                in a medium with bubbling air containing 5% carbon
                dioxide"
BASE COUNT      74 a      153 c      148 g      75 t
ORIGIN
Query Match      6.1%; Score 38.6; DB 10; Length 450;
Best Local Similarity 50.8%; Pred. No. 9.6;
Matches 92; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 183 CTTGGCTTTACTGTCGTCTGACCAATTCACGTTCCGCTTATGAGTGCGCAACGTGTC 242
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 221 CTTGACACCCCTGGACGGCTGTCATCGTCGACACCAACTTCGCGTGGGACACATCTT 280

QY 243 CGGATGTACCATGTACAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACGCGA 302
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 281 CGGCGTGTGGCCCGCGACAGCTGACCAACATCACCAGCGCGACGAGTGCTGCCGG 340

QY 303 CATGATCATGCACACCCCGGCTGCTCCCTGCTGGGAGAACAACTTCCGCTG 362
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 341 CATGGCATCTACGGTCCCGCACCGCTGTCTGATTGCTTGAAGGACGCCCGGCTG 400

QY 363 C 363

```

Db	401	C	401
RESULT 11			
LOCUS	AV637328	451 bp	mRNA
DEFINITION	AV637328 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii		
ACCESSION	AV637328		
VERSION	AV637328.1	GI:10780648	
KEYWORDS	EST.		
SOURCE	Chlamydomonas reinhardtii.		
ORGANISM	Chlamydomonas reinhardtii.		
REFERENCE	Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;		
AUTHORS	Chlamydomonadaceae; Chlamydomonas.		
TITLE	1 (bases 1 to 451)		
JOURNAL	Asamizu, E., Miura, K., Kuchio, K., Inoue, Y., Fukuzawa, H., Ohyama, K.,		
MEDLINE	Nakamura, Y., and Tabata, S.		
COMMENT	Generation of expressed sequence tags from low-CO2 and high-CO2		
	adapted cells of Chlamydomonas reinhardtii		
	DNA Res. 7 (5), 305-307 (2000)		
	20539644		
	Contact: Erika Asamizu		
	The First Laboratory for Plant Gene Research		
	Kazusa DNA Research Institute		
	Yama 1532-3, Kisarazu, Chiba 292-0812, Japan		
	Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.		
FEATURES	Location/Qualifiers		
source	1..451		
	/organism="Chlamydomonas reinhardtii"		
	/db_xref="taxon:3055"		
	/clone="HC071f12_r"		
	/clone_1db="Chlamydomonas reinhardtii 5% CO2"		
	/note="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2:		
	XhoI; The cDNA library was constructed from cells cultured		
	in a medium with bubbling air containing 5% carbon		
	dioxide"		
BASE COUNT	74 a 150 c 150 g 77 t		
ORIGIN			
Query Match	6.1%; Score 38.6; DB 10; Length 451;		
Best Local Similarity	50.8%; Pred. No. 9.6;		
Matches	92; Conservative 0; Mismatches 89; Indels 0; Gaps 0;		
Oy	183 CTTGGCTTTACTGTCCTCTGTGACCAATTCACAGCTTCGCTTATGAGTGGCGACAGTGTG	242	
Db	178 GTTCACACCCCTGGAGCGGCTGTCATCGTGACACCAACTTCGCGGTGGGACCATCTT	237	
Oy	243 CGGGATGTACCACTGCACGAAGCACTGCTCCAACTGCACAGCAATGTGTATGAGCGACGGA	302	
Db	238 CGGCGTGTGGCCCGCGGACGACAGCTGCACCACTCACCGCGCGGACAGTGGCTGCGG	297	
Oy	303 CATGATCATGCACACCCCGGGGTGGTCCCTCGCTTCGCGGAGAACACTTCTCCGCTG	362	
Db	298 CATGGGCAATCAAGCGTCCCGGACGCGTGTCTGCATTGCGCCCTGGAAGGACGCGCGGCTG	357	
Oy	363 C 363		
Db	358 C 358		
RESULT 12			
LOCUS	AV637643	451 bp	mRNA
DEFINITION	AV637643 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii		
ACCESSION	AV637643		
VERSION	AV637643.1	GI:10780963	
KEYWORDS	EST.		
SOURCE	Chlamydomonas reinhardtii.		
ORGANISM	Chlamydomonas reinhardtii		

[illegible]

	Query Match	6.1%	Score 38.6;	DB 10;	Length 456;	
	Best Local Similarity	50.8%;	Pred. No.9.6;			
Matches	92;	Conservative	0;	Mismatches	89;	Indels 0; Gaps 0;
OY	183 CTTGGCTTATACGTGTCCTTGCTGCAGCAATTCCACCATTCCCGTTATGTAGAGTGCACAAGCTGTC					242
Dd	111 111 111 111 111 111 111 111 111 111 111 111 111 111 111					111
OY	52 CTTCGAACCCCCCTGGAGAGGCTGCTCCTCATGCTGCACAACTTCGCCGTGGGACCATCTT					111
OY	243 CGGAGTAGTAACCAATGTCAGCAAGACTGCTCCAACTCAATCAAGCAATTTGTATAGAGCACGGA					302
Dd	112 CGGCGGTGTGGCCCCGGAGAACGTGACCAAACATCACACCGGCCCGAGAGGTGGTCCGG					171
OY	303 CATGATCATAGCAACACCCC GGSGGTGGCCCGTGCAGTGGAGAACAATCTTCCCGTG					362
Dd	172 CATGGGCATCTACGAGTCCCGCCGACCGTGTTCGTGATTTGCTGTGAAGAGCCGCCGGCTG					231
OY	363 C 363					
Dd	232 C 232					
RESULT 15						
LOCUS	AV635382	456 bp	mRNA	linear	EST 15-DEC-2000	
DEFINITION	AV635382 Chlamydomonas reinhardtii 5' CO2 Chlamydomonas reinhardtii					
ACCESSION	CDNA clone HC045110_r 5', mRNA sequence.					
VERSION	AV635382					
KEYWORDS	AV635382.1 GI:10778702					
SOURCE	EST.					
ORGANISM	Chlamydomonas reinhardtii. Chlamydomonas reinhardtii Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales; Chlamydomonadales; Chlamydomonas.					
REFERENCE	1 (bases 1 to 456) Asamizu,Y., Miura,K., Kuchio,K., Inoue,Y., Fukuzawa,H., Ohyaama,K., Nakamura.Y. and Tabata,S. Generation of expressed sequence tags from low-CO2 and high-CO2 adapted cells of Chlamydomonas reinhardtii					
AUTHORS	DNA Res. 7 (5), 305-307 (2000)					
TITLE						
JOURNAL MEDLINE	20539644					
COMMENT	Contact: Erika Asamizu The First Laboratory for Plant Gene Research Kazusa DNA Research Institute Yana 1532-3, Kisarazu, Chiba 292-0812, Japan Email: asamizuekazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/. Location/Qualifiers					
FEATURES	source					
	1..456					
	/organism="Chlamydomonas reinhardtii"					
	/strain="C9"					
	/db_xref="taxon:3055"					
	/clone="HC045f10_r"					
	/clone_lib="Chlamydomonas reinhardtii 5% CO2"					
	/note="Vector: pBluescriptII SK-; Site_1: EcoRI; site_2: XhoI; The cDNA library was constructed from cells cultured in a medium with bubbling air containing 5% carbon dioxide"					
BASE COUNT	75 a 158 c 148 g 75 t					
ORIGIN						
* Query Match	6.1%; Score 38.6; DB 10; Length 456;					
Best Local Similarity	50.8%; Pred. No.9.6;					
Matches	92; Conservative	0; Mismatches	89; Indels	0; Gaps	0; Gaps	0;
OY	183 CTTGGCTTATACGTGTCCTTGCTGCAGCAATTCCACCATTCCCGTTATGTAGAGTGCACAAGCTGTC					242
Dd	221 CTTCGAACCCCCCTGGAGAGGCTGCTCCTCATGCTGCACAACTTCGCCGTGGGACCATCTT					280
OY	243 CGGAGTAGTAACCAATGTCAGCAAGACTGCTCCAACTCAATCAAGCAATTTGTATAGAGCACGGA					302
Dd	281 CGGCGGTGTGGCCCCGGAGAACGTGACCAAACATCACACCGGCCCGAGAGGTGGTCCGG					340
OY	303 CATGATCATAGCAACACCCC GGSGGTGGCCCGTGCAGTGGAGAACAATCTTCCCGTG					362

Db 341 CATGGGCACTCTACGGTCCCGCACCGCTGTTCTGCAATTGCCCTGAGAGACGCCCCGCGCTG 400
Qy 363 C 363
Db 401 C 401

Search completed: February 18, 2003, 07:07:42
Job time : 1698 secs